Characterization of equine metabolic syndrome and mapping of candidate genetic loci

A DISSERTATION
SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL
OF THE UNIVERSITY OF MINNESOTA
BY

Nichol Elizabeth Schultz

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
Doctor of Philosophy

Molly E. McCue, Advisor
James R. Mickelson, Co-Advisor

January 2016

Acknowledgements

First, I would like to express gratitude to my brilliant and supportive advisor, Dr. Molly McCue. Thank you for providing me the skills necessary to develop into an independent researcher in such an exciting, challenging, and rewarding field of study.

Second, I would to thank my co-advisor, Dr. Mickelson. Thank you for sharing your scientific knowledge, experience, and superb editing skills.

Next, I would like to acknowledge my committee, Drs. Geor, Martinson, and Reed. Thank you for your guidance and support. I appreciate your efforts to keep me on task while completing this dissertation.

Throughout my studies, I was fortunate to have received funding from multiple sources including the Pfizer-Morris Animal Health Research Fellowship and the American Association of Equine Practitioners Past-Presidents Research Fellowship. I am sincerely grateful for their financial support.

Past and present lab members also deserve acknowledgement. Thank you Jessica Petersen, Krista Fritz, Annette McCoy, Raffaella Teixeira, Rob Schaefer, Michelle Lucio, Katie Minor, Kari Ekenstedt, Eva Furrow, Leah Freilich, Elaine Norton, and Sam Beeson for your help and friendship throughout the years.

Dedication

This dissertation is dedicated to my husband, Tom, who has been a constant source of encouragement and support from the beginning of this journey through the end.

Abstract

Equine metabolic syndrome (EMS) is a clustering of clinical signs associated with increased risk of laminitis, a potentially life-threatening condition of the foot. Similar to human metabolic syndrome (MetS), generalized and/or regional adiposity, hyperinsulinemia, insulin resistance and dyslipidemia, are reported components of EMS. However, there is ongoing debate regarding the definition of EMS, its etiology and pathogenesis, and the mechanisms linking EMS to its secondary consequences. Conflicting reports regarding EMS reflect the limitations of prior EMS studies, and that EMS is likely a complex, multifactorial condition similar to MetS. The primary objectives of this thesis were to characterize metabolic variation and EMS across horse and pony breeds and to identify candidate genes for EMS risk.

Chapter 2 details the largest-ever epidemiological investigation of EMS in which 11 metabolic traits were measured in >600 horses and ponies from 166 farms. The use of multivariate, multilevel regression modeling allowed, for the first time, quantification of the relative importance of environmental (farm, dietary composition, exercise, etc.) and individual (age, breed, sex etc.) factors on these metabolic traits, while accounting for the often strong correlation between the trait measures. Age, sex, breed, obesity, prior laminitis status, and time of year were all strongly associated with one or more metabolic traits. Despite strong associations, these factors only explained 9.6% to 36.3% of the variation across these 11 traits, thus the majority of the variability in these measures remained unexplained. Unexplained variation at the farm level after accounting for diet, exercise, and sampling time of year, suggests that additional unmeasured environmental factors explain the similarity in metabolic measures between horses sampled from the same farm. Similarly, unexplained variation at the individual level suggests that unmeasured individual characteristics, for example genetics, are responsible for a large proportion of individual trait variation.

Differences in the incretin response may also contribute to individual trait variation. The incretin response, defined as the difference in insulinemic responses between an oral and intravenous glucose challenge, is controlled by intestinal secretion of peptides, such as GLP-1, that stimulate pancreatic insulin secretion. While the incretin response has

been hypothesized to play a role in the EMS pathogenesis, this hypothesis has not been adequately tested. In Chapter 3, the glycemic, insulinemic, and total and active GLP-1 responses to an oral sugar challenge, and the activity of DPP4, the major protease that breaks down GLP-1, were characterized. The use of a longitudinal analysis, rather than the traditional area under the curve analysis, allowed for more power to detect differences in these responses, including variation due to breed, obesity, and prior laminitis status.

Unexplained individual level variation and breed differences in metabolic phenotypes support the hypothesis that there is an underlying genetic susceptibility to EMS. The final objective of this thesis was to identify candidate genes associated with EMS. MetS is a highly polygenic syndrome where numerous candidate genes have been identified. Whereas MetS associated variants are typically of small effect size; it was hypothesized that in EMS a small number of moderate to large effect loci contribute to variation in metabolic traits due to the fact that horse populations do not randomly mate and experience substantial selection pressure. 286 Morgan horses were genotyped on the Illumina SNP50 chip and imputed up to $>$800,000 SNPs to perform a genome wide association study (GWAS) to identify candidate genes for EMS. Additive genetic variance estimated from a genomic relationship matrix calculated from genotyped SNPs ("chip heritability") indicated that the 11 measured metabolic traits were moderately heritable. Yet initial genome-wide scans using standard linear mixed models failed to detect significant associations.

In Chapter 4, an improved linear mixed model for mapping polygenic traits in a population with familial relationships similar to that in many equine GWAS was developed and validated. The model incorporates a Bayesian variable selection method to rank SNPs and a stepwise feature selection process to determine the optimal SNPs to model the random polygenic effect, while including a random effect for each sampled herd or "familial cluster". The method was validated using the QTL-MAS 2010 dataset, and Morgan horse and Welsh pony height datasets, and demonstrated increased power while controlling the false positive rate.

Using this improved linear mixed model, 76 suggestive and 17 genome-wide significant candidate loci were identified for the 11 metabolic traits in the 286 Morgan horse cohort. Candidate genes had a substantial overlap with MetS candidate genes such as

VEGFA, *NRXN3*, *GRIK2*, and *TRIB2*. Other interesting candidate genes included *ISL*, which encodes insulin enhancer protein that is thought to play an important role in regulating insulin gene expression; and *AHR* which encodes the aryl hydrocarbon receptor, a ligand activated transcription factor known to bind endocrine disrupting chemicals such as polycyclic aromatic hydrocarbons and dioxins. *AHR* is an interesting candidate gene given the potential role of endocrine disrupting chemical in the pathophysiology of MetS, and unexplained sources of farm level variation in Chapter 2.

A unifying theme of Chapters 2-5 was the similarities between EMS and MetS, and the complex phenotypic and genetic architecture in both species. The use of advanced statistical modeling approaches allowed for a more complete understanding of the metabolic phenotypic variation in Chapters 2 and 3, and for the identification of many associated genetic loci in Chapter 5. The shared candidate genes for metabolic syndrome in humans and horses suggests similar underlying pathophysiological mechanisms and provides opportunity for exploring similar preventative and therapeutic management strategies.

Contents

List of Tables

List of Figures

List of Abbreviations

AACE	American Association of Clinical Endocrinologists
ACTH	adrenocorticotropin hormone
ACVIM	American College of Veterinary Internal Medicine
ADAD1	adenosine deaminase domain containing 1
AHR	aryl hydrocarbon receptor
AIC	Akaike Information Criterion
ALK	anaplastic lymphoma receptor tyrosine kinase
APN	adiponectin
APOA5	apolipoprotein A-V
APOC3	apolipoprotein C-III
ATG14	autophagy related 14
AUC	area under the curve
BCS	body condition score
BMI	body mass index
BRINP3	bone morphogenetic protein/retinoic acid inducible neural-specific 3
BSLMM	bayesian sparse linear mixed model
BVSR	bayesian variable selection regression
bwt	bodyweight
CCDC54	coiled-coil domain containing 54
CCSER2	coiled-coil serine rich protein 2
CETP	cholesteryl ester transfer protein
CNST	consortin, connexin sorting protein
CNTNAP4	contactin associated protein-like 4
COBL	cordon-bleu WH2 repeat protein

COMMD9	COMM domain containing 9
CP	crude protein
CREB5	cAMP responsive element binding protein 5
CRP	C-reactive protein
CRYBA4	crystallin beta A4
CTNNA3	catenin alpha 3
CTPS1	CTP synthase 1
CVD	cardiovascular disease
DACT3	dishevelled-binding antagonist of beta-catenin 3
DAPK2	death-associated protein kinase 2
df	degrees of freedom
DIO3	deiodinase, iodothyronine, type III
DIP2B	disco interacting protein 2 homolog B
DLEU7	deleted in lymphocytic leukemia, 7
DLGAP4	discs, large (Drosophila) homolog-associated protein 4
DM	dry matter
DNA	deoxyribonucleic acid
DPP4	dipeptidyl peptidase-4
ECA	Equus caballus
EGIR	European Group for the study of Insulin Resistance
EM	expectation-maximization
EPB41L4B	erythrocyte membrane protein band 4.1 like 4B
ET-1	endothelin-1
FAM155A	family with sequence similarity 155 member A
FAM50B	family with sequence similarity 50 member B
FLVCR2	feline leukemia virus subgroup C cellular receptor family member 2
FRK	fyn-related Src family tyrosine kinase
FSIGTT	frequently-sampled intravenous glucose tolerance test
FTO	Fat mass and obesity-associated protein
GATB	glutamyl-tRNA(Gln) amidotransferase, subunit B
GBA3	glucosidase, beta, acid 3 (gene/pseudogene)
GCG	proglucagon gene (*GCG*)

GH	girth to height ratio
GIP	glucose-dependent insulinotropic polypepide
GLP-1	glucagon-like peptide-1
GLP1a	active GLP-1
GLP1tot	total GLP-1
GLU	glucose
GNGT1	guanine nucleotide binding protein (G protein), gamma transducing activity polypeptide 1
GPR75-ASB3	GPR75-ASB3 readthrough
GPRIN3	GPRIN family member 3
GRAMD1C	GRAM domain containing 1C
GRIK2	glutamate receptor, ionotropic, kainate 2
GRM	genomic relationship matrix
GUE	proglucagon gene (*GCG*) upstream enhancer region
HDAC1	histone deacetylase 1
HDL	high-density lipoproteins
HIVEP3	human immunodeficiency virus type I enhancer binding protein 3
hMetS	human metabolic syndrome
HS3ST3B1	heparan sulfate-glucosamine 3-sulfotransferase 3B1
ICC	intra-cluster correlation coefficient
IDF	International Diabetes Federation
IL-6	interleukin-6
INS	insulin
ISL1	ISL LIM homeobox 1
JPH1	junctophilin 1
KB	kilobase
KCNT2	potassium channel, sodium activated subfamily T, member 2
KCTD18	potassium channel tetramerization domain containing 18
KIF2A	kinesin heavy chain member 2A
KLHDC1	kelch domain containing 1
KLRC1	killer cell lectin-like receptor subfamily C, member 1
LAM	laminitis

LD	linkage disequilibrium
LDL	low density lipoprotein
LEP	leptin
LGI1	leucine-rich, glioma inactivated 1
LMM	linear mixed model
LYPD4	LY6/PLAUR domain containing 4
MAF	minor allele frequency
MAP	mitogen activated protein
MAP10	microtubule associated protein 10
MB	megabase
Mcal	megacalories
MCMC	Markov chain Monte Carlo
MDGA2	MAM domain containing glycosylphosphatidylinositol anchor 2
MEGF10	multiple EGF-like-domains 10
MetS	metabolic syndrome
MFN1	mitofusin 1
MPZL2	myelin protein zero-like 2
NCEP/ATP	National Cholesterol Education Program Adult Treatment Panel
NCOA1	nuclear receptor coactivator 1
NDF	neutral detergent fiber
NH	neck circumference to height ratio
NOVA1	neuro-oncological ventral antigen 1
NRXN3	neurexin 3
NS	Nichol Schultz
NUDT12	nudix hydrolase 12
ob	obese
OMA1	OMA1 zinc metallopeptidase
OR10K1	olfactory receptor family 10 subfamily K member 1
OR2H1	olfactory receptor family 2 subfamily H member 1
OTOL1	otolin 1
PAI-1	plasminogen activator inhibitor-1
PCDH18	protocadherin 18

PCDH7	protocadherin 7
PCR	polymerase chain reaction
PI3K	phosphoinositide 3-kinase
PNLIPRP1	pancreatic lipase-related protein 1
PPID	pars pituitary intermedia dysfunction
QTL	quantitative trait loci
QTL-MAS	QTL mapping and marker assisted selection
RAB8B	RAB8B, member RAS oncogene family
REML	restricted maximum likelihood
RFU	relative fluorescence units
RHOBTB2	Rho-related BTB domain containing 2
RIMS1	regulating synaptic membrane exocytosis 1
RIPK1	receptor interacting serine/threonine kinase 1
RN7SKP34	RNA, 7SK small nuclear pseudogene 34
RNF217	ring finger protein 217
RPL22L1	ribosomal protein L22 like 1
SCNN1B	sodium channel, non voltage gated 1 beta subunit
SE	standard error
SEL1L2	sel-1 suppressor of lin-12-like 2 (C. elegans)
SENP5	SUMO1/sentrin specific peptidase 5
SEPP1	selenoprotein P, plasma, 1
SLC23A2	solute carrier family 23 (ascorbic acid transporter), member 2
SLCO4C1	solute carrier organic anion transporter family member 4C1
SNORD12	small nucleolar RNA, C/D box 12
SNP	single nucleotide polymorphism
SOX5	SRY-box 5
SOX6	SRY-box 6
SPAG16	sperm associated antigen 16
ST3GAL1	ST3 beta-galactoside alpha-2,3-sialyltransferase 1
ST6GALNAC2	ST6 (alpha-N-acetyl-neuraminyl-2,3-beta-galactosyl-1,3)-N-acetylgalactosaminide alpha-2,6-sialyltransferase 2
t-PA	tissue-type plasminogen activator

T2D	type 2 diabetes
TCF7L2	Transcription Factor 7-Like 2
TG	trigylceride
TNFα	tumor necrosis factor alpha
TNS3	tensin 3
TRBV19	T cell receptor beta variable 19
TRIB2	tribbles pseudokinase 2
TRPS1	trichorhinophalangeal syndrome I
TSN	translin
TTC29	tetratricopeptide repeat domain 29
USP25	ubiquitin specific peptidase 25
UTR	untranslated region
VEGFA	vascular endothelial growth factor A
VLDL	very low density lipoproteins
WC	waist circumference
WHO	world health organization
WSC	water soluble carbohydrates
XPO4	exportin 4
ZFP69B	ZFP69 zinc finger protein B
ZKSCAN2	zinc finger with KRAB and SCAN domains 2
ZNF362	zinc finger protein 362

Chapter 1

Introduction and Literature Review

1.1 Introduction

Metabolic syndrome, a constellation of abnormalities, including obesity, dyslipidemia, glucose intolerance and hypertension, has become a global health concern largely due to the increase in prevalence and association with increased risk of cardiovascular disease and type 2 diabetes (T2D). A similar metabolic condition exists in horses and is associated with increased risk of laminitis, a crippling and potentially life-threatening condition of the foot. Ongoing debate exists in both the human and veterinary fields regarding the etiology and pathogenesis of metabolic syndrome and the mechanisms linking the syndrome to its secondary consequences. In both fields, conflicting reports generated from different investigations of the etiology and pathogenesis of metabolic syndrome reflect the complex, multifactorial nature of the condition. In order to develop effective preventative and therapeutic strategies for managing the secondary consequences of the syndrome, further research is warranted to determine the mechanisms underlying metabolic syndrome.

1.2 Review of metabolic syndrome literature

1.2.1 Human metabolic syndrome

Definition of human metabolic syndrome

Metabolic syndrome (MetS) is a growing public-health challenge worldwide coinciding with urbanization, excess dietary intake, increasing obesity and sedentary lifestyles. MetS itself does not cause clinical signs but rather comprises a set of risk factors for diseases that include atherosclerosis, coronary heart disease, stroke, and T2D. MetS began as a concept rather than a diagnosis and has been evolving for years. In 1920, Kylin demonstrated the association of hypertension, hyperglycemia, and gout.[1] In 1947, Vague described an association of visceral adiposity with metabolic abnormalities found with cardiovascular disease (CVD) and type II diabetes.[2] At the 1965 meeting of the European Association for the Study of Diabetes, a syndrome was reported to be comprised of hypertension, hyperglycemia, and obesity.[3] In 1988 Raeven described a cluster of risk factors for T2D and CVD and suggested the name "Syndrome X". He introduced the concept of insulin resistance and compensatory hyperinsulinemia as the main pathophysiological mechanism underlying MetS.[4] Many names have been given to the syndrome although they all appear to describe different components of the same entity. Recognized components of MetS include insulin resistance, obesity, dyslipidemia, hypertension, and hyperglycemia.

Establishment of MetS as a disease entity has been impeded by the lack of a consensus regarding criteria for its diagnosis. Several groups including the world health organization (WHO),[5] European Group for the study of Insulin Resistance (EGIR),[6] the American Association of Clinical Endocrinologists (AACE),[7] the National Cholesterol Education Program Adult Treatment Panel (NCEP/ATP)[8]and International Diabetes Federation (IDF) have attempted to set forth diagnostic criteria for MetS based on cut-off points for different MetS component measurements.[9,10] A major challenge in developing criteria that is applicable worldwide is to define obesity cut-off points. Different populations have different distributions of norms for body weight and weight circumferences and the relationship between these values and the risk of CVD or T2D varies in different populations. Because no single definition of metabolic syndrome has been accepted worldwide, data regarding the prevalence of MetS is variable. Worldwide

prevalence ranges from $< 10\%$ to as much as $> 80\%$ depending on the regional environment, population composition, and the definition of MetS used.[11,12] Overall, the IDF estimates 25% of the world's adult population has MetS.[9]

Pathophysiology of human metabolic syndrome

The etiology and pathogenesis of MetS is an area of active debate with many unknowns. To quote Zimmet on the IDF's stance "MetS may be an etiologic mystery but far from myth".[13] Insulin resistance and obesity are the dominant factors cited in the development of MetS. The current discussion of MetS pathophysiology is largely based on the recognition of adipose tissue as a physiologically active organ along with the recognition of the concept of insulin resistance and its consequences.

The role of adipose tissue in MetS pathophysiology:
Adipose tissue was once considered as merely a storage site for lipids but is now recognized as a complex, active endocrine tissue that secretes many factors referred to as adipokines that regulate metabolism and vascular biology. Examples of adipokines include: adiponectin, leptin, resistin, and pro-inflammatory mediators (tumor necrosis factor alpha (TNFα) and interleukin-6 (IL-6)), plasminogen activator inhibitor-1 (PAI-1), and C-reactive protein (CRP).[14]

Adiponectin regulates lipid and glucose metabolism, increases insulin sensitivity, regulates food intake and bodyweight, and has anti-inflammatory effects. Adiponectin inhibits endothelial activation, reduces conversion of macrophages to foam cells, and inhibits smooth muscle proliferation characteristic in the development of atherosclerosis.[15] Low adiponectin is a strong independent risk factor for CVD and is inversely correlated with blood pressure, low-density lipoprotein cholesterol, triglycerides, insulin resistance, and hyperinsulinemia independent of fat mass.[16–18]

Leptin secretion is regulated by adipose tissue mass. Plasma leptin levels increase during the development of obesity and decline with weight loss. Leptin receptor signaling in the hypothalamus and brain stem regulates satiety and energy intake.[19] However, obese individuals exhibit leptin resistance and do not experience appetite suppression in the presence of elevated leptin levels as expected. Leptin resistance is thought to play a key role in the pathology of obesity.[14] In addition, leptin affects blood pressure

via activation of the sympathetic nervous system.[20] Hyperleptinemia is considered an independent risk factor for CVD development.[21]

Resistin expression is linked with insulin resistance and is 15-fold greater in rodent visceral adipose tissue than subcutaneous adipose tissue.[22] Serum resistin is elevated with rodent obesity and treatment with recombinant resistin in rodents produces insulin resistance.[23] However, human resistin shares limited homology with murine resistin and is expressed at low levels in human adipocytes. Numerous epidemiological studies in humans have failed to provide a consistent link between resistin expression in adipose tissue or circulating levels with adiposity or insulin resistance.[22]

Adipocyte hypertrophy and hyperplasia occurs in response to nutrient excess. As the adipocytes enlarge, blood supply is reduced and hypoxic damage to the adipocytes occurs.[24] Necrosis and macrophage infiltration of the adipose tissue leads to the production of pro-inflammatory cytokines, which include pro-inflammatory mediators TNFα and IL-6, PAI-1, and CRP.[19] Circulating TNFα is secreted primarily by macrophages in obese adipose tissue, whereas adipocytes typically produce a membrane bound form.[25,26] TNFα stimulates lipolysis and increased levels are highly correlated with insulin resistance, hyperinsulinemia, and hypertension.[27–29] TNFα also stimulates expression of other inflammatory mediators such as leptin and IL-6 and inhibits expression and secretion of adiponectin.[30–32] IL-6 has been shown to inhibit insulin action in muscle, liver, and adipocytes and has been reported to be an important contributor to the chronic inflammatory state and hepatic insulin resistance of obesity.[33,34] PAI-1 is a serine protease inhibitor secreted from adipocytes, platelets, and vascular endothelium that inhibits tissue plasminogen activator and is therefore a marker of impaired fibrinolysis and atherothrombosis.[19,35] PAI-1 levels are increased in individuals with abdominal obesity.[36] CRP levels are positively correlated with waist circumference, insulin resistance, body mass index (BMI), and hyperglycemia.[37–39] CRP is also more likely to be elevated in obese insulin resistant individuals than non-obese insulin resistant individuals.[40]

In summary, adipocytokines affect insulin signaling, energy metabolism, blood coagulation, and inflammatory responses demonstrating adipose tissue does much more than store and mobilize lipids.

Insulin Resistance:

Insulin resistance is defined as a pathophysiological condition in which a normal insulin concentration does not produce a normal insulin response in the target tissue such as muscle, liver, or adipose. Pancreatic β-cells secrete more insulin (hyperinsulinemia) to compensate to some degree in order to maintain normoglycemia (T2D is the result of β-cell exhaustion hence inadequate compensating for insulin resistance). However, under hyperinsulinemic conditions, increased insulin levels at target tissues results in the clinical manifestations of MetS via differential activation of insulin signaling pathways. Insulin signaling occurs following binding of insulin to the insulin receptor and tyrosine phosphorylation of downstream substrates and activation of two parallel pathways: the phosphoinositide 3-kinase (PI3K) pathway and the mitogen activated protein (MAP) kinase pathway. The PI3K pathway is inhibited under insulin resistant conditions while the MAP kinase pathway functions normally. Inhibition of the PI3K pathway leads to a reduction in endothelial nitrous oxide production causing endothelial dysfunction along with a reduction in GLUT4 translocation resulting in decreased glucose uptake in skeletal muscle. Since the MAP kinase pathway is still functioning normally, there is continued endothelin-1(ET-1) production, expression of vascular cell adhesion molecules, and stimulation of vascular smooth muscle cells leading to the vascular abnormalities predisposing to atherosclerosis.[41] Thus, continued functioning of the MAP kinase insulin signaling pathway combined with inhibition of the PI3K pathway under conditions of insulin resistance resulting in endothelial dysfunction is a possible mechanism linking MetS to atherosclerosis.

Obesity enhances insulin resistance. In obese individuals, adipose tissue releases increased amounts of non-esterified fatty acids (NEFA), glycerol, hormones, pro-inflammatory cytokines and other factors involved in the development of insulin resistance.[42] NEFA are derived mainly from adipose tissue triglyceride stores released through the action of hormone sensitive lipase. NEFA are also produced through lipolysis of lipoproteins in tissues by the action of lipoprotein lipase.[43] Insulin is involved in both anti-lipolysis and the stimulation of lipoprotein lipase. The most sensitive pathway of insulin action is inhibition of lipolysis in adipose tissue. However, when insulin resistance develops, the increased lipolysis of stored triacylglycerol molecules in adipose tissue produces more

free fatty acids, further inhibiting the anti-lipolytic effect of insulin and creating additional lipolysis.[44] In humans, insulin resistance develops within hours of an acute increase in NEFA levels.[45] Moreover, insulin mediated glucose uptake improves with an acute decrease in NEFA levels following treatment with an anti-lipolytic agent.[46] Acute exposure of skeletal muscle to elevated NEFA levels induces insulin resistance by inhibiting insulin-mediated glucose uptake and secondly, chronic exposure of the pancreas to elevated NEFA levels impairs β-cell function.[47]

Dyslipidemia:

Insulin resistance leads to an atherogenic dyslipidemia in several ways. First, as described above, impaired insulin signaling impairs the normal suppression of lipolysis by insulin resulting in increased NEFA levels. The NEFA flux to the liver and serve as a substrate for triglyceride synthesis. NEFA also stabilize the production of apoB, the major lipoprotein of very low-density lipoproteins (VLDL) particles, resulting in increased VLDL production. Secondly, since insulin normally degrades apoB through PI3K-dependent pathways, insulin resistance directly increases VLDL production due to impairment of the PI3K pathway. Third, insulin regulates the activity of lipoprotein lipase, which is a major mediator of VLDL clearance. Overall, insulin resistance results in hypertriglyceridemia via both an increase in VLDL production and a decreased clearance of VLDL. VLDL is metabolized to small, dense LDL which promote atheroma formation.[48] In summary, the NEFA flux to the liver of insulin resistant individuals is increased resulting in increased triglyceride synthesis and storage, and excess triglyceride is secreted as VLDL. It is thought that the dyslipidemia associated with insulin resistance is a direct consequence of increased VLDL secretion by the liver.[49–51]

Hypertension:

Two mechanisms have been proposed for the hypertension that occurs in insulin resistant individuals. Both insulin and leptin activate the sympathetic nervous system resulting in increased vascular tone and hypertension. Secondly, insulin resistance is known to activate the renin-angiotensin-aldosterone system (RAAS). Renin facilitates conversion of angiotensinogen to angiotensis I that is subsequently converted to angiotensin II by

angiotensin-converting enzyme. Angiotensis II causes blood vessels to constrict resulting in increase blood pressure. Aldosterone increases blood pressure mainly through sodium retention and plasma volume expansion. RAAS may also promote oxidative stress and endothelial dysfunction.[52]

Prothrombotic state and endothelial dysfunction:
PAI-1 increases with insulin resistance. PAI-1 inhibits tissue-type plasminogen activator (t-PA), a serine protease that converts plasminogen to plasmin and is involved in clot specific fibrinolysis. Thus increased PAI-1 contributes to a prothrombotic state.[53] Endothelial dysfunction is characterized by an impaired endothelium-dependent vasodilation. Hyperinsulinemia causes the release of ET-1, a potent vasoconstrictor. Whereas adiponectin stimulates the production of nitric oxide (anti-inflammatory vasodilator) in vascular endothelial cells,[54] suggesting adiponectin has a protective vascular effect and that this protective effect is weakened in individuals with low adiponectin levels.[55] The potent vasoprotective effects of nitric oxide counteract various atherogenic processes such as vascular smooth muscle cell proliferation, platelet adhesion and thrombogenesis, lipid peroxidation, and monocyte adhesion to endothelial cells.[56]

In summary, resistance to the actions of insulin and compensatory hyperinsulinemia play a central role in MetS, as either the cause or consequence of one or more of its components.

Genetic susceptibility to human metabolic syndrome

Large variation in MetS susceptibility in individuals with a similar risk profile suggests an interaction between environmental and genetic factors. For example, some individuals who are not obese are insulin resistant where as some individuals who are obese are not insulin resistant.[57] Ethnic variation exists in the pattern of metabolic risk factors. Expression of metabolic risk factors may be under genetic control, which influences the response in different types of environments. The "thrifty genotype hypothesis" proposed by Neel hypothesizes individuals living in harsh environments with a fluctuating food supply would increase their probability of survival if they were able to maximize storage of surplus energy. Genetic selection would favor energy-conserving genotypes in these

environments. However, this genetic adaptation may become unfavorable in alternative environments.[58]

Family studies suggest a genetic basis for the components of metabolic syndrome. Significant genetic correlations among the MetS components have been identified for BMI, waist circumference, high-density lipoproteins (HDL), triglycerides, insulin, and PAI-1.[59] Candidate gene studies and genome wide association studies have revealed numerous candidate genes for MetS. In a systematic review, Povel describes the most studied single nucleotide polymorphisms (SNPs) in relation to MetS suggesting an association of MetS with SNPs in the *FTO*, *TCF7L2*, *IL6*, *APOA5*, *APOC3*, and *CETP* genes.[60] The *FTO* polymorphism is a top GWAS hit for BMI[61] and the *TCF7L2* polymorphism is a top hit for T2D.[62] The *IL6* polymorphism is associated with increased BMI and IL-6 levels.[63] *APOA5*, *APOC3*, and *CETP* polymorphisms are all associated with hypertriglyceridemia.[64–66]

1.2.2 Animal models of human metabolic syndrome

For a complex disease syndrome like MetS no single, ideal animal model can be expected to meet all of the research needs. Therefore a large number of different species and strains are used to study causative and pathophysiologic hypotheses related of MetS.

Rodent models are used to test different MetS causative hypotheses such as genetics, fetal programming, diet, activity, aging, gender, and pollution. The ability to introduce or eliminate genes from the rodent genomes allows testing of single-gene mutations and the polygenic basis of MetS. Quantitative trait loci (QTL) mapping conducted in rodents is used to locate new candidate genes of MetS traits that often become targets of human candidate gene studies.[67] Rodent models are used to determine how and when altered early nutrition and growth affect adult development of MetS.[68] Diet-induced rodent models are used to examine the role of diet in MetS. The two most commonly studied models are the high-sucrose fed spontaneously hypertensive rats and the high-fructose fed Sprague Dawley rats (SD) which develop insulin resistance, hyperinsulinemia, and hypertension.[69] Diet-induced obese rats and mice are commonly used to study the role of obesity in MetS, although variation in glucose tolerance, insulin resistance, and triglyceride levels depend on the strain, gender, and source of dietary fat.[70] Classic obesity rodent models also include single-gene loss-of-function mutations related to leptin

metabolism that result in extreme obesity such as obese (ob/ob) mice (Lep^{ob} mutation) and the Zucker (fa/fa) obese rat ($Lepr^{fa}$).[71]

Other animal species such as sheep, pigs, rabbits, and dogs have also been used to study the etiology and pathogenesis of MetS. The correlation between visceral adiposity and insulin resistance in dogs closely resembles that of humans.[72] Rodents and rabbits do not spontaneously replicate atherosclerosis seen in humans partly due to differences in lipoprotein metabolism which include lack of cholesterylester transfer protein, differences in the apolipoprotein B pathway, and VLDL dominated lipid metabolism in rabbits. Cholesterol feeding and mechanical endothelial injury are used to induce atherosclerosis in these models.[73] Pigs have LDL-dominated lipid metabolism similar to humans and do develop atherosclerosis.[74]

1.2.3 Equine metabolic syndrome

Current definition of equine metabolic syndrome

A condition similar to human metabolic syndrome exists in horses, referred to as equine metabolic syndrome (EMS), that was first described by Johnson in 2002[75] and accepted by a consensus committee in 2010.[76] The equine condition resembles the human condition in many aspects, although the vascular consequences differ. In humans, the vascular structures affected by the condition are typically the coronary vessels whereas in horses the vasculature of the hoof is affected resulting in laminitis, a potentially severe, crippling and life-threatening condition. Johnson recognized the primary features of the laminitis-prone phenotype shared striking similarities to those described for human MetS including obesity and insulin resistance. The 2010 American College of Veterinary Internal Medicine (ACVIM) listed several criteria for EMS based on the research data available at the time. The three main criteria included:[76]

- Documented or suspected *insulin resistance*, ie hyperinsulinemia and/or abnormal glycemic and insulinemic responses to oral or IV glucose or insulin challenges;
- *Generalized obesity* and/or increased adiposity in specific locations (*regional adiposity*) including the nuchal ligament ("cresty neck"), the tail head, behind the shoulder, in the prepuce or mammary gland region;

- A predisposition toward *laminitis* that develops in the absence of other recognized causes, such as grain overload, retained placenta, colitis, colic or pleuropneumonia.

Additional suggested components included:

- Hypertriglyceridemia, dyslipidemia, and increased low density lipoprotein concentrations[77–79]
- Hypertleptinemia[80]
- Arterial hypertension[78,81]
- Altered reproductive cycling in mares[82,83]
- Increased systemic markers of inflammation in association with obesity[84]

Features of equine metabolic syndrome

Similar to the debate and evolving definition of human MetS, the features that define EMS are a subject of ongoing debate in the equine veterinary community despite the earlier consensus statement. Descriptions of the metabolic phenotype of laminitis-prone horses and ponies have varied among published studies (Table 1.1),[77–79,81,85] making a unifying phenotypic definition difficult. The lack of consensus among study reports may result from unmeasured explanatory variables, insufficient sample size to detect significant effects in the presence of confounding variables, and differences in experimental design including differences in test cohort, breed, time of sample collection (time of day and or season). Lack of consensus also likely reflects the complexity of the "EMS" phenotype with multiple factors at both individual and environmental levels likely contributing to variation in metabolic traits.

Insulin resistance is the primary criteria used by clinicians and investigators to define EMS. However, true insulin resistance is rarely documented in EMS cases. The diagnosis of insulin resistance in horses is complicated by the fact that "gold standard" methods for assessing insulin sensitivity (frequently-sampled intravenous glucose tolerance test, hyperinsulinemic euglycemic clamps) are labor intensive and not readily used in clinical practice. Further, the substitution methods for diagnosing insulin resistance such as fasting insulin levels, combined glucose-insulin tolerance tests, and oral glucose

Table 1.1: A summary of findings related to obesity, regional adiposity and endocrine/metabolic variables in published studies of the equine metabolic syndrome (EMS) phenotype.

	Treiber et al.[77] 2006*	Frank et al.[79] 2006	Bailey et al.[81] 2008	Carter et al.[78] 2009*
Breed(s)	Welsh and Dartmoor ponies	6 breeds	Mixed-breed ponies	Welsh and Dartmoor ponies
Sample size	160	12	80	74
Obesity (BCS)	Yes	Yes	No	Yes
Regional adiposity	Yes	Yes	No	Yes
Hyperinsulinemia	Yes	Yes	Yes**	Yes
Insulin resistance	Yes(RISQI)	Yes(CGIT)	Yes**(RISQI)	Yes(RISQI)
Fasting glucose	Not different	Higher in EMS	Not different	Not different
Triglycerides	Higher in EMS	Not different	Higher in EMS**	Higher in EMS
NEFAs	Not different	Higher in EMS	Not evaluated	Not evaluated

*Data obtained from the same population of Welsh and Dartmoor ponies; **serum insulin and triglyceride concentrations and RISQI differed between ponies with and without a history of laminitis in summer but not in winter; BCS = body condition score; RISQI = reciprocal of the square root of the serum insulin concentration; CGIT = combined glucose-insulin tolerance test; NEFAs = nonesterified fatty acids; Table adapted from McCue et al.[85]

tolerance tests are highly variable and have not been thoroughly evaluated as valid measures of insulin sensitivity. Thus, while there has been much discussion about insulin resistance in EMS, very few studies have reported quantitative data on insulin sensitivity and other aspects of glucose and insulin dynamics in affected animals. Two studies using minimal model analysis of a frequently-sampled intravenous glucose tolerance test (FSIGTT) have provided evidence of compensated insulin resistance in laminitis predisposed breeds.[86,87]

The majority of reports have used "fasting" or "resting" measures of insulin and glucose and/or indices derived from minimal model analysis of FSIGTT measurements as proxy indicators of insulin resistance,[88] with the current consensus view that a fasting insulin concentration >20 mIU/L indicates insulin resistance.[76] Fasting hyperinsulinemia is typically (but not always) accompanied by normoglycemia, which suggests compensated insulin resistance, i.e. increased pancreatic insulin secretion in response to reduced tissue insulin sensitivity. Two studies in ponies using minimal model analysis of FSIGTT have provided evidence of compensated insulin resistance, based on indices of insulin sensitivity and acute insulin response.[86,89] Mechanisms other than a compensatory increase in pancreatic secretion may contribute to the hyperinsulinemia. Decreased insulin clearance by the liver is a second possible mechanism of hyperinsulinemia. In one study, reduced insulin clearance was shown to contribute to hyperinsulinemia in obese horses.[90] It has also been proposed that hyperinsulinemia may not necessarily be secondary to insulin resistance but that the reverse may actually be true,[91] leading some to question whether hyperinsulinemia in EMS is the "cart or the horse". It has been considered plausible that horses manifesting chronic hyperinsulinemia may be the result of an enhanced incretin response (intestinal secretion of incretin hormones that stimulate insulin secretion in response to oral sugar) with insulin resistance occurring secondary to chronic hyperinsulinemia.[92,93] Hyperinsulinemia has been shown to be a feature in the EMS phenotype in ponies and Morgan horses,[77–79,81] although there is wide variation in values among studies. The prevalence of hyperinsulinemia (insulin concentration > 20 mIU/L) was 18% of the obese/over-conditioned horses vs only 1.4% of the normally conditioned horses in a study of 300 randomly selected mixed-breed horses in southwest Virginia,[94] while another study in Australia reported a 27% prevalence of hyperinsulinemia in randomly selected ponies.[95] In the pony study, age, body

condition score (BCS), supplementary feeding and a history of laminitis were identified as risk factors for hyperinsulinemia.

Evidence of obesity and/or the presence of one or more enlarged subcutaneous fat deposits (regional adiposity) has been regarded as one of the defining characteristics of EMS. Horses and ponies diagnosed with EMS may be obese, defined as a Henneke scales BCS of $\geq$7/9, and/or have regional adiposity. Common sites of regional fat accumulation include the nuchal ligament region (cresty neck), behind the shoulder (uni- or bilateral), around the tail-head, and in the preputial or mammary gland regions. Similar to human MetS, generalized obesity and/or regional adiposity are thought to play a role in the EMS phenotype. Studies in horses have demonstrated an inverse relationship between BCS and specific measures of insulin sensitivity and positive correlations between adiposity, resting insulin concentrations and blood markers of inflammation (e.g. serum amyloid A, TNF-α, mRNA encoding for IL-1$\beta$and TNF-α).[78,84,96,97] Induction of an ~20% weight gain in Arabian geldings has been shown to result in a reduction in insulin sensitivity,[98] while weight loss in obese ponies resulted in improvement in glucose tolerance and decreased fasting insulin concentrations.[99] Interestingly, moderate weight gain in Thoroughbred geldings failed to alter insulin sensitivity.[100] These observations reinforce the idea that factors other than BCS and adipose tissue mass contribute to differences in insulin sensitivity, but also raise the possibility that there are breed differences in regards to the impact of obesity on metabolic function.

Increased plasma triglyceride concentrations were a feature of the EMS phenotype in a closed herd of Welsh ponies[77] and was also documented in an out-bred population of ponies with a history of recurrent laminitis,[81] although hypertriglyceridemia was variable with season. In a small, mixed-breed group of obese, insulin-resistant horses increased serum NEFA concentrations (but not triglycerides) were detected, whereas serum NEFAs were not useful in the differentiation of an EMS phenotype in ponies.[77,79]

Hypertension, altered serum adipokine concentrations, and systemic inflammation, all of which are features of metabolic syndrome in humans, may be additional components of the EMS phenotype. Bailey and colleagues[81] detected arterial hypertension in mixed-breed, recurrent laminitic ponies during summer but not in winter. Increased serum leptin has been reported in Welsh ponies with other features of the EMS phenotype,[78] while other studies have reported a positive association between BCS and

circulating inflammatory biomarkers[84,96] and a significant inverse relationship between apparent adiposity and serum adiponectin.[101,102] Decreased adiponectin levels have been detected in ponies with prior laminitis history.[103] A study of 15 ponies in a weight reduction program demonstrated a link between plasma adiponectin and insulin resistance, whereas serum leptin was linked to adiposity, independent of insulin sensitivity.[104]

Features of EMS seem to be more common is some breeds compared to others, especially pony breeds in comparison to horse breeds. Other breed suggested predispositions include the Morgan breed, Miniature horses, Spanish Mustang, Saddlebred, Warmblood, Haflinger, Norwegian Fjord, Peruvian Paso, Paso Fino breeds, and Tennessee Walking horses.[105] EMS has also been reported in some Quarter Horses. Familial clustering of EMS associated laminitis has also been identified in pony breeding lines supporting a genetic basis for the condition in horses.[77]

Laminitis has been demonstrated to be inducible by infusing super physiological amounts of insulin intravenously over 2-3 day period.[106,107] The mechanism linking insulin and laminitis is unknown with hypotheses stemming from human MetS and believed links between insulin, insulin resistance, hypertension, and endothelial dysfunction. A possible mechanism being impairment of the PI3K pathway resulting in decreased nitric oxide synthesis with subsequent reduced digital vasodilator effects, activation of the MAPK pathway stimulating ET-1 synthesis and sympathetic nervous system activation with subsequent digital vasoconstriction effects.[76]

The horse as a potential animal model of metabolic syndrome

As described above, significant clinical overlap exists between features of human and equine metabolic syndrome (Table 1.2).[85] Both conditions place the individual at a higher risk of a vascular condition, atherosclerosis in humans and laminitis in horses. The underlying etiology for both conditions are still unknown and subject of current debate in both the human and veterinary medical fields. It is possible these two similar conditions may share a common etiology and pathogenesis and knowledge from one species may be transferable to the other. As stated earlier, no one animal can be expected to serve all research needs. Equine models of MetS have the potential to provide novel insight into the underlying etiology and pathophysiologic mechanisms of human MetS and its sequelae.

Table 1.2: Human and equine metabolic syndrome comparison

Human Metabolic Syndrome (MetS)	Equine Metabolic Syndrome (EMS)
Increased BMI or obesity	Generalized obesity assessed by body condition score (BCS)
Intraabdominal or visceral obesity	Regional adiposity in the nuchal ligament "cresty neck" phenotype strongly correlated with serum glucose, insulin, triglyceride and leptin concentrations, and glucose tolerance
Hyperinsulinemia	Hyperinsulinemia fasting insulin concentrations >20 mU/L
Impaired fasting glucose	Resting plasma glucose concentrations increased in comparison with normal horses in some studies
Insulin resistance	Insulin resistance suggested by various oral and IV glucose challenge models
Hypertension	Hypertension
Dyslipidemia: increased NEFA, triglycerides, decreased VLDL, HDL	Increase in triglycerides, nonesterified fatty acids (NEFA), HDL cholesterol, decreased VLDL
Proinflammatory state: increased TNF-α, IL-1, IL-6	Increased TNF-α, IL-1, decreased IL-6
Altered adipokines: increased leptin, decreased adiponectin	Increased leptin, decreased adiponectin

Abbreviation: BMI, body mass index; HDL, high-density lipoprotein; IL, interleukin; TNF-α, tumor necrosis factor alpha; VLDL, very-low-density lipoprotein; Table adapted from McCue et al.[85]

1.3 Review of multilevel modeling: basics and extensions

Careful consideration must be given to experimental design and analytical approaches directed at understanding the complex, multifactorial basis of metabolic syndrome. As described in the above literature review, conflicting results have been reported for both human and equine metabolic syndrome studies, adding to the confusion regarding the definition and pathophysiology of metabolic syndrome. The following section is intended to provide a brief overview of an analytical approach utilized in this thesis to address the analytical challenges presented when studying a complex, multifactorial condition with

numerous correlated features impacted by both individual and environmental factors.

Table 1.3: Multilevel modeling definitions.

Multilevel model statistical term	Short explaination
Random effect	Levels of the factor are a only a "sample" of the total possible levels.
Fixed effect	Levels of the factor are not a "sample" of possible levels, i.e. all levels are sampled.
Intracluster correlation coefficient (ICC)	A measure of homogeneity of sampled clusters that can be interpreted two ways: 1) proportion of total variance of an outcome accounted for by the clusters or 2) the correlation between measured outcomes for two randomly drawn individuals in the cluster.
"Null" model	Model does not contain explanatory variables, only decomposes the variance of an outcome.
"Full" model	Model contains predictor variables that explain variance in an outcome.
R^2	Proportion of variance explained by predictor variables.

1.3.1 Reasons for multilevel models

A multilevel statistical model is applied to stratified data in order to elucidate relationships at more than one level.[108] A multilevel linear model (aka hierarchical linear model,[109] random coefficients model,[110] mixed effects model,[111] growth curve model[112]) is a standard method for analyzing outcome data with hierarchical structure such as observations clustered within individuals and individuals clustered within groups. The hierarchical structure leads to correlations among the observations within a cluster and therefore violation of the statistical assumptions of outcome independence. Violating this assumption leads to inaccurate standard errors and significance tests for fixed effects. Multilevel models account for the correlation among observations by including additional error terms (random effects).

Epidemiological investigations often feature a clustered sample design. Modeling the

outcome dependence is important to control for bias as described above but can also be of substantive interest. Multilevel models can assess the influence of cluster level explanatory variables while controlling for individual level explanatory variables. This thesis describes an epidemiological investigation of metabolic trait variation in horses sampled with a clustered design; multiple horses sharing a common farm environment were sampled from numerous farms throughout the United States. Sources of metabolic trait variation likely include both individual and environmental factors. Examples of individual factors include genetics, age, and sex where as examples of environmental factors include diet and season. A multilevel analysis facilitates an investigation of metabolic syndrome at both an individual and environmental level.

1.3.2 The basic two-level multilevel model

The hierarchical linear model notation writes a separate equation for each level of the hierarchy and then combines the equation into a single model equation. The first level of the model for simple regression on the level 1 variable X is written as

$$\text{Level 1: } Y_{ij} = \beta_{0j} + \beta_1 x_{1ij} + e_{0ij}, \tag{1.1}$$

where i (1,...,n) refers to the level 1 (individual) unit and j (1,...,m) to the level 2 (group) unit. The level 1 model appears similar to a typical ordinary least squares multiple regression model, however, the j subscripts indicate a different level 1 model is being estimated for each of the j level 2 groups. The intercept and/or slopes are allowed to randomly vary across level 2 units. An initial step toward modeling between-group variability is to let the intercept vary between groups. In the case of a random intercept model each of the different j groups are allowed to have a mean that differs than the population mean. The level 2 equation is written as

$$\text{Level 2 random intercept: } \beta_{0j} = \gamma_{00} + u_{0j}. \tag{1.2}$$

The notation for the regression coefficients is changed and the average intercept is called γ_{00} while the regression coefficient for X is called γ_{10}. Substitution leads to the combined

model

$$Y_{ij} = \underbrace{\gamma_{00} + \gamma_{10} x_{ij}}_{\text{fixed}} + \underbrace{u_{0j} + e_{0ij}}_{\text{random}}, \tag{1.3}$$

where $E(u_{0ij}) = 0$ and $var(u_{0ij}) = \sigma^2_{u0}$ and $E(e_{0ij}) = 0$ and $var(e_{0ij}) = \sigma^2_{e0}$. The single model equation indicates which part of the model comprises the fixed effects (the γs) and which part comprises the random effects (u and e). The single model equation illustrates that the level 1 parameters (β_{0j}, β_{1j},β_{2j}) are not directly estimated but are indirectly estimated through the level 2 gammas γs.[108]

For example, in a sample population with multiple sampled farms and sampling of multiple horses per farm the horse is considered level 1 and the farm is considered level 2. A random intercept effect is estimated for each farm, i.e. the difference of the trait mean for the farm from the trait mean of the entire sample population.

Intracluster correlation coefficient (ICC)

The intracluster correlation coefficient (ICC) measures the proportion of variation accounted for by the group level. The sum of the level 2 and level 1 variances reasonably estimates the sample variance of the dependent variable.[113] The parameter is called a correlation coefficient because it is equal to the correlation between values of two randomly drawn individuals from the same group where (σ^2_{u0} and σ^2_{e0} are estimates of the level 2 and level 1 variances obtained by fitting a null model (empty model with only an intercept and no explanatory variables)).[108]

$$\begin{gathered} \text{Null (“empty”) model: } Y_{ij} = \gamma_{00} + u_{0j} + e_{0ij} \\ \text{ICC: } \rho = \frac{\sigma^2_{u0}}{(\sigma^2_{u0} + \sigma^2_{e0})} \end{gathered} \tag{1.4}$$

An ICC greater than zero indicates the observations are not independent and justifies multilevel modeling of the data. Multilevel modeling relaxes the assumption of independence and allows for correlated error structures. If ordinary least squares is used for clustered data with correlated errors, the resulting standard errors are biased downward resulting in a greater chance of Type 1 errors. Multilevel models estimate the

appropriate unbiased errors. Therefore it was important to perform a multilevel model analysis of the data in this study given horses were not sampled randomly and therefore can not be considered independent observations, an assumption of standard ordinary least squares regression. If independent observations were assumed, test statistics may be biased and lead to incorrect inferences.

Assessing model fit

Models fit by restricted maximum likelihood (REML) that are nested and only differ in their random effects can by assessed for their goodness of fit with the REML statistic. The Akaike Information Criterion (AIC=-2 loglikelihood + 2*number or parameters, decreases with goodness of fit) and the loglikelihood (increases with goodness of fit) can be used for assessment of relative differences in goodness of fit[114] and in the case of model comparisons, the chi-square distributed likelihood ratio and its associated p-value.[108]

In ordinary least square regression the fit of a model is assessed by calculating R^2 the percentage of variance of the outcome accounted for by explanatory variables in the model. In multilevel models there is a separate R^2 for each level of the model and it is possible to have smaller or even negative R^2 values with inclusion of additional explanatory variables in the model. Alternative R^2 approaches for multilevel models often referred to as a pseudo-R^2. The statistic is interpreted as the proportional reduction in variance or proportion of variance explained for a parameter estimate that results from the use of one model as compared to a Null model or alternative base model with fewer explanatory variables.[109]

$$\begin{aligned} \text{pseudo-R}^2_{level\ 2} &= \frac{\text{Null model } \sigma^2_{u0} - \text{Full model } \sigma^2_{u0}}{\text{Null model } \sigma^2_{u0}} \\ \text{pseudo-R}^2_{level\ 1} &= \frac{\text{Null model } \sigma^2_{e0} - \text{Full model } \sigma^2_{e0}}{\text{Null model } \sigma^2_{e0}} \end{aligned} \tag{1.5}$$

1.3.3 Extending the basic multilevel model

Longitudinal models

Multilevel modeling can also be applied to longitudinal or repeated measures data where multiple observations are clustered within a single individual or object and the primary interest is in modeling the predictors of change over time. Multilevel models are able to characterize the between and within-individual variability of an outcome trajectory over time and identify factors that influence the trajectory.

Initially, plots of the data over time should be examined to determine the appropriateness of a linear, quadratic, or cubic relationship with time. An ordinary least squares regression model fit to an unconditional (aka "Null","empty") model with time included as the only predictor aids in determining whether a linear, quadratic, or cubic modeling of time best fits the mean population trajectory. Second, multilevel models are fit sequentially to determine the appropriateness of allowing the intercept, linear slope, or quadratic slope to vary randomly between the level 2 individuals (AIC and likelihood ratio tests aid in the selection of the random regression model). For example, consider a quadratic model of time was determined to best model the mean population trajectory. The level 1 equation is written

$$\text{Unconditional Level 1: } Y_{ij} = \beta_{0j} + \beta_{1j}(\text{time})_{ij} + \beta_{2j}(\text{time}^2)_{ij} + e_{0ij}, \tag{1.6}$$

where i (1,...,n) refers to the level 1 unit (time point) and j (1,...,m) to the level 2 unit (individual). The level 2 equations modeling a random intercept, linear slope, and quadratic slope are written

$$\begin{aligned} \text{Unconditional Level 2: } \beta_{0j} &= \gamma_{00} + u_{0j} \\ \beta_{1j} &= \gamma_{10} + u_{1j} \\ \beta_{2j} &= \gamma_{20} + u_{2j}, \end{aligned} \tag{1.7}$$

where $E(u_{0ij}) = E(u_{1ij}) = E(u_{2ij}) = 0$, $var(u_{0ij}) = \sigma^2_{u0}$, $var(u_{1ij}) = \sigma^2_{u1}$, $var(u_{2ij}) = \sigma^2_{u2}$, $cov(\sigma^2_{u0}, \sigma^2_{u1}) = \sigma_{u01}$, $cov(\sigma^2_{u0}, \sigma^2_{u2}) = \sigma_{u02}$, $cov(\sigma^2_{u1}, \sigma^2_{u2}) = \sigma_{u12}$, $E(e_{0ij}) = 0$ and

$var(e_{0ij}) = \sigma^2_{e0}$. The combined unconditional model is written as

$$\text{Combined unconditional: } Y_{ij} = \gamma_{00} + \gamma_{10}(\text{time})_{ij} + \gamma_{20}(\text{time}^2)_{ij} + u_{0j} + u_{1j} + u_{2j} + e_{0ij}. \tag{1.8}$$

The next step is to move on to a conditional model where explanatory variables considered to influence the intercept or slopes are included to the model. It is important to note that the interpretation of the intercept and slopes are dependent on the coding of time. The intercept is the mean value when time=0 and the linear slope is the rate of change at time=0. For a quadratic model, the quadratic slope remains the same at all time points. The quadratic slope is equivalent to one half of the acceleration/deceleration of the rate of change. Below are examples of level 1 and level 2 equations for a conditional model that includes the effect of level 2 explanatory variable x_j on the intercept, linear slope and quadratic slope.

$$\begin{aligned}
\text{Conditional Level 1: } & Y_{ij} = \beta_{0j} + \beta_{1j}(\text{time})_{ij} + \beta_{2j}(\text{time}^2)_{ij} + e_{0ij} \\
\text{Conditional Level 2: } & \beta_{0j} = \gamma_{00} + \gamma_{01}x_j + u_{0j} \\
& \beta_{1j} = \gamma_{10} + \gamma_{11}x_j + u_{1j} \\
& \beta_{2j} = \gamma_{20} + \gamma_{21}x_j + u_{2j} \\
\text{Combined equation: } & Y_{ij} = \gamma_{00} + \gamma_{01}x_j + \gamma_{10}(\text{time})_{ij} + \gamma_{11}x_j(\text{time})_{ij} \\
& \quad + \gamma_{20}(\text{time}^2)_{ij} + \gamma_{21}x_j(\text{time}^2)_{ij} \\
& \quad + u_{0j} + u_{1j} + u_{2j} + e_{0ij}
\end{aligned} \tag{1.9}$$

A longitudinal model is used in Chapter 3 to model glucose, insulin and GLP-1 measurements repeated at 7 timepoints (level 1) for each horse (level 2).

Three level models

Multilevel models can be extended to handle more than two levels. Repeated measurements are often collected from individuals that are clustered in groups or alternatively individuals may be sampled from subgroups of a larger group. For example, if the repeated observations for individuals modeled in equation (1.8) were correlated at a higher level due to being sampled in a clustered design the model can be extended to account for correlation at the third level. The mean for each group is allowed to vary

randomly across level 3 groups. The level 1, level 2, level 3, and combined model for repeated measures clustered within individuals clustered within groups where the mean is allowed to vary randomly across level 3 units and individuals at level 2 are allowed to have randomly varying intercepts and slopes would be written as

$$\begin{aligned}
\text{Unconditional Level 1:} \quad & Y_{ijk} = \beta_{0jk} + \beta_{1jk}(\text{time})_{ijk} + \beta_{2jk}(\text{time}^2)_{ijk} + e_{0ijk} \\
\text{Unconditional Level 2:} \quad & \beta_{0jk} = \delta_{00k} + u_{0jk} \\
& \beta_{1jk} = \gamma_{100} + u_{1jk} \\
& \beta_{2jk} = \gamma_{200} + u_{2jk} \\
\text{Unconditional Level 3:} \quad & \delta_{00k} = \gamma_{000} + v_{00k} \\
\text{Combined unconditional:} \quad & Y_{ijk} = \gamma_{000} + \gamma_{100}(\text{time})_{ijk} + \gamma_{200}(\text{time}^2)_{ijk} \\
& + v_{00k} + u_{0jk} + u_{1jk} + u_{2jk} + e_{0ijk}
\end{aligned} \tag{1.10}$$

where i (1,...,n) refers to the level 1 unit (time point), j (1,...,m) to the level 2 unit (individual), and k (1,...,l) to the level 3 unit (group).

A three level model is used in Chapter 3 to model glucose, insulin and GLP-1 measurements repeated at 7 timepoints (level 1) for each horse (level 2) sampled from a particular farm (level 3).

Multivariate multilevel models

Multilevel models can be extended to handle more than one dependent outcome. The dependent variable Y_{hij} is the measurement on the hth variable for individual i in group j. It is possible to analyze all h variables independently although there are several reasons for considering analyzing the data in a multivariate manner.[113] For one, partitioning of the covariances between the dependent variables over the different levels of the model allow conclusions to be drawn about the correlations between the dependent variables at both group and individual levels. Second, the tests of specific effects for a single dependent variable will be more powerful in the multivariate analysis in the situation where individuals are missing data for some of the dependent variables but the dependent variables are strongly correlated. Third, a multivariate approach allows one to test whether the effect of an explanatory variable on dependent variable

Y_1 is larger than its effect on Y_2 when Y_1 and Y_2 were observed on the same individuals. Finally, a multivariate approach is needed if one desires to carry out a single test of the joint effect of an explanatory variable on several dependent variables.

The random intercept model for dependent variable Y_h with individual-dependent or group-dependent explanatory variables X_1,...,X_p is written as

$$Y_{hij} = \gamma_{h00} + \gamma_{h10}x_{1ij} + \gamma_{h20}x_{1ij} + ... + \gamma_{hp0}x_{pij} + u_{h0j} + e_{0hij}. \tag{1.11}$$

Three nesting levels are used to represent multivariate data in a multilevel manner. The first level is the dependent variables indexed by h=1,...,m, the second level is the individuals i=1,...,n_j, and the third level is the groups j=1,...,N. Each measurement of a dependent variable far a particular individual is represented by a single line of the data matrix containing Y_{hij}, h, i, j, and explanatory variables x_1,...,x_p. Dummy variables d_1,...,d_m are used to indicate the dependent variable represented on the data line.

$$d_{shij} = \begin{cases} 1 & (h = s), \\ 0 & (h \neq s) \end{cases} \tag{1.12}$$

The random intercept model for the m dependent variables is combined into one three level hierarchical linear model. In the sums over s=1,...,m only the term for s=h contributes and all the other terms disappear.

$$Y_{hij} = \sum_{s=1}^{m} \gamma_{0s} d_{shij} + \sum_{k=1}^{p} \sum_{s=1}^{m} \gamma_{ks} d_{shij} x_{kij} + \sum_{s=1}^{m} u_{sj} d_{shij} + \sum_{s=1}^{m} e_{0sij} d_{shij} \tag{1.13}$$

Since the dependent variables Y_1,.... Y_m are measured on the same individuals their dependence can be taken into account. The (u)s and the (e)s are components of vectors

$$u_{ij} = \begin{Bmatrix} u_{10j} \\ \vdots \\ u_{m0j} \end{Bmatrix}, \quad e_{0ij} = \begin{Bmatrix} e_{01ij} \\ \vdots \\ e_{0mij} \end{Bmatrix}. \tag{1.14}$$

There are now residual covariances matrices present at the group and individual level instead of residual variances.

$$\underbrace{\begin{bmatrix} \sigma^2_{u1} & \sigma_{u1u2} & \dots & \sigma_{u1um} \\ \sigma_{u1u2} & \sigma^2_{u2} & \dots & \sigma_{u2um} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{u1um} & \sigma_{u2um} & \dots & \sigma^2_{um} \end{bmatrix}}_{\text{between group covariance matrix}} \quad \underbrace{\begin{bmatrix} \sigma^2_{e01} & \sigma_{e01e02} & \dots & \sigma_{e01e0m} \\ \sigma_{e01e02} & \sigma^2_{e02} & \dots & \sigma_{e02e0m} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{e01e0m} & \sigma_{e02e0m} & \dots & \sigma^2_{e0m} \end{bmatrix}}_{\text{within group covariance matrix}} \tag{1.15}$$

A multivariate model is used in Chapters 2 to model the covariance structure of metabolic phenotypes at both the individual horse and farm level. The multivariate model is used in Chapter 3 to model the covariance structure of trajectories at individual horse level.

1.4 Hypothesis and objectives

The underlying molecular mechanism responsible for EMS has not been identified. It is clear however, that metabolic syndrome itself and its key components, obesity, insulin resistance and laminitis are influenced by environmental factors, with EMS most commonly seen in horses receiving excessive nutrition, often from pasture. On the other hand, nutritional excess does not result in EMS in all horses. The factors underlying this difference in susceptibility have not been determined, but recent studies indicate it may be due to an underlying genetic predisposition. Currently, the genetic variants underlying this disease phenotype are unknown, which restricts understanding of the disease pathophysiology, limits the ability to predict disease risk, and hinders the ability to identify individuals who can benefit from early intervention.

Central hypothesis: An underlying genetic susceptibility to EMS exists and is shared across horse breeds. The goal is to define the EMS phenotype and shared phenotypic risk factors across breeds and to identify the underlying genetic component/risk alleles.

Objective 1: Quantify the variation in metabolic phenotypes across horse/pony breeds and the impact of individual and environmental factors, and to identify differences in metabolic phenotypes in horses/ponies with obesity and/or a history of laminitis.

Previous studies of the EMS phenotype have demonstrated variability in phenotypic measurement across equine populations yet a large scale, across-breed epidemiologic investigation of EMS has not been performed. In this aim, epidemiologic data including signalment/history, environmental (diet and physical activity) and phenotypic data (morphometric and biochemical measures) will be collected from 5 target breeds and a total of 600+individuals. Data will be analyzed to:

1. Describe the correlations between morphometric, biochemical, and hormonal measures used to assess equine metabolic variation.

2. Quantify the proportion of metabolic trait variation between farms and individual horses within farms.

3. To estimate the effect of measured individual factors (i.e. age, breed, and gender) and measured environmental factors (i.e. diet, exercise, management practices, etc.) on metabolic trait measures.

4. To determine how metabolic trait measurements differ between obese and non-obese horses with and without previous laminitis.

5. Quantify the proportion of explained by environmental and individual factors by quantifying the variation explained and unexplained variation at the farm and individual horse level in metabolic trait measurements.

6. Determine the relative importance of age, breed, sex, obesity, prior laminitis status, diet, and exercise to the explained variability in metabolic traits.

Objective 2: To examine response to an oral sugar challenge and incretin biology (DPP-IV activity; insulin secretory and GLP-1 responses; SNPs in GCG and DPP4) in Morgan horses and Welsh ponies. The hypothesis is that DPPIV activity and GLP-1 responses to oral glucose challenge differ significantly between EMS horses and unaffected animals, and that these differences in incretin responses are associated with SNPs in DPP4 and GCG. Objectives include:

1. Characterize DPP-IV activity, insulin secretory and GLP-1 responses to oral glucose challenge in Morgan horses and Welsh ponies; and

2. Determine if SNPs within GCG and DPP4 are associated with differences in insulin responses, DPP-IV activity and GLP-1 responses.

Objective 3: Identify genetic loci associated with equine metabolic trait variation.

1. Develop and validate an improved linear mixed for mapping polygenic traits in a population with familial relatedness.

2. Perform a genome-wide association analysis in single breed cohort (genotyped using the Equine SNP50 Beadchip) with EMS associated phenotypes identified in Objective 1.

Significance of studying equine metabolic syndrome

In summary, both human and equine metabolic syndrome are increasingly common conditions with devastating consequences. Knowledge and understanding of the features of metabolic syndrome in both species and their link to vascular consequences are increasing, however much still remains unknown. The ongoing debates evolving in both human and veterinary fields regarding the etiology and pathogenesis of metabolic syndrome reflect the complex and likely multifactorial basis of the condition. Further research directed at characterizing the EMS phenotype and genetic basis of the condition is warranted and will benefit horses and potentially humans. Determination of the genetic risk factors underlying the development of EMS and laminitis susceptibility will lead to an improved understanding of EMS pathophysiology. This will enable identification of horses at risk for EMS and allow preventative management practices to be put into place and also aid in the identification of therapeutic strategies.

Chapter 2

Re-defining the equine metabolic syndrome phenotype

2.1 Summary

Chapter 2 details the largest-ever epidemiological investigation of equine metabolic syndrome (EMS), a clinical syndrome in horses associated with increased risk of laminitis development. 11 metabolic traits were measured in 610 horses and ponies from 166 farms. The use of multivariate, multilevel regression modeling allowed, for the first time, quantification of the relative importance of environmental (farm, dietary composition, exercise, etc.) and individual (age, breed, sex etc.) factors on these metabolic traits, while accounting for the often strong correlation between the trait measures. Age, sex, breed, obesity, prior laminitis status, and time of year were all strongly associated with one or more metabolic traits. Despite strong associations, these factors only explained 9.6% to 36.3% of the variation in these traits, thus the majority of the variability in these measures remains unexplained. Unexplained variation at the farm level after accounting for diet, exercise, and time of year, suggests that additional environmental factors explain the similarity in metabolic traits measured from horses sharing the same farm environment. Similarly, unexplained variation at the individual level suggests unmeasured individual characteristics, for example genetics, may be responsible for metabolic trait variation.

2.2 Introduction

Equine metabolic syndrome (EMS) is a clinical syndrome associated with increased risk of laminitis development. Generalized/regional adiposity, hyperinsulinemia, insulin resistance, dyslipidemia, and predisposition to laminitis development are reported components of the syndrome. However, not all studies have consistently reported the same components for EMS (see Table 1.1). Hyperinsulinemia has been shown to be a feature of the EMS phenotype in ponies[77,81] and Morgan horses,[79] however, values vary between studies.[78,81,115] Elevated plasma triglyceride concentrations have been associated with the EMS phenotype in ponies[78,81,116] but not in horses;[79] whereas, non-esterified fatty acid (NEFA) concentrations have been associated with an EMS phenotype in horses,[79] but were not useful in the differentiation of an EMS phenotype in ponies.[77,78] Morphometric measurements to quantify generalized obesity (girth to height ratio and body condition score) and regional adiposity (neck circumference to height ratio) have also been associated with the EMS phenotype in horses[79] and proposed as predictors of laminitis risk in ponies;[78] but obesity is not a consistent finding across studies, particularly in well-managed populations.[81]

The body of knowledge concerning EMS is growing, however there is a lack of consensus regarding the EMS phenotypic criteria. Potential explanations include differential in the experimental design between studies. For example, differences in the study cohort or time of year of sample collection result in different conclusion due to breed or season. Insufficient sample size or unmeasured confounding variables may be an additional limiting factor of previous studies. In addition, EMS is not likely a singular condition; multiple factors at both individual and environmental levels likely contribute to variation in metabolic traits.

The objectives of this chapter are to quantify the variation in metabolic phenotypes across horse/pony breeds and the impact of individual and environmental factors, and to identify differences in metabolic phenotype in horses/ponies with obesity and/or a history of laminitis. To address these objectives, a cross-sectional study of 11 metabolic traits was performed in 610 horses clustered within 166 farms. A multilevel, multivariate statistical model was utilized to quantify the impact of both individual and farm characteristics on metabolic measurements in this cohort (trait variation) and determine

how these measured responses relate to each other (trait co-variation). In a multivariate model, multiple outcome measures are analyzed simultaneously. A multilevel model partitions the variance of the outcomes into both individual (horse) and group (farm) levels. By modeling a multivariate, multilevel outcome response one recognizes that the metabolic traits measured in this cohort are inter-related and that EMS is a complex condition, measured by several outcome variables (metabolic phenotypes). And further, that these outcomes are impacted by both individual and environmental factors. Analyzing data in which samples are clustered by farm environment using ordinary linear regression is problematic due to violation of the assumptions of independent observations. Increased similarity of observations from individuals within the same farm leads to underestimation of the standard errors for regression parameter estimates and inflates the Type I error.[117] Multilevel modeling is well suited to this type of data structure and is able to model the outcome of interest by examining how farm-level or horse-level characteristics are related to metabolic trait measures.

When evaluating multiple outcome measurements, one approach would be to perform separate multilevel analyses for each trait. However, this option can lead to an inflation of type 1 error (finding a significant effect where in reality there is no effect), especially when the measured traits are inter-related as with EMS phenotypes. An alternative approach is to model a multivariate, multilevel model. A multivariate model has several advantages over a series of univariate analyses. By accounting for the possible relationships between different outcome variables in a multivariate model, one decreases the chance of making a type 1 error. For example, in a situation where two outcome measurements are correlated and an explanatory variable effects only one of the outcome variables; there is a good chance univariate analysis would find an association with each outcome because of their correlation structure, despite one of the associations being untrue. A second advantage of a multivariate, multilevel model over a series of univariate models is that multivariate models typically have greater statistical power. Finally, the multivariate approach enables us to assess the degree of correlation between different metabolic traits and the extent to which correlations depend on individual and environmental factors.

2.3 Methods

2.3.1 Study Population

Horses were recruited into the study using two mechanisms to ensure an adequate representation of horses with varying metabolic phenotypes from different breeds and shared environments. A 2-step online data survey system was designed to recruit owners of EMS suspect horses. The initial survey obtained information regarding the horses' signalment, previous history of laminitis, and propensity toward being an "easy keeper". Exclusion criteria for further participation in the study included suspicion of pars pituitary intermedia dysfunction (PPID) based on age, hair coat, increased thirst/urination, and/or diagnostic test results for PPID. Inclusion criteria to participate in a second, more detailed survey included previous history of laminitis (not associated with systemic infection or non-weight bearing lameness in contralateral limb), obesity, and/or being considered an "easy keeper". The second survey obtained information from the EMS suspect horse and along with information from a second horse on the property not suspected of EMS. Information obtained included digital photos of the horses, body measurements, diet/exercise information, frequency/treatment of laminitis, and signalment information from the non-EMS suspect horse. Based on survey responses, horse owners were invited to submit blood and dietary samples from their horses. In addition to the online recruitment of study participant, Morgan and Welsh pony breed associations were also contacted to identify farm owners with 1 or more EMS suspect horses who were willing to allow researchers to visit the farm and sample and obtain information from multiple horses.

Samples were obtained from 634 horses, 17 horses were eliminated from the study population due to evidence of hair coat abnormalities and not already defined adrenocorticotropin hormone (ACTH) elevation (>90pg/ml) suggestive of PPID. Seven horses were eliminated from the study population due to incomplete data (failure to submit diet samples, unknown prior laminitis status). The mean study population age was used for 1 horse with unknown age. A total of 610 horses were included in the analysis.

2.3.2 Sampling Protocol

Horses were removed from pasture the evening prior to morning (7-10 am) blood sample collection performed by either the researchers or the horse owner's veterinarian. Horses were allowed access to water and provided with 1 flake of hay at 10PM. Horse owners were also given the opportunity to have the horse participate in an optional oral sugar test. Following collection of the baseline blood sample, 0.15 ml corn syrup per kg bodyweight was administered orally and a second blood sample obtained 75 minutes post oral sugar administration. Blood samples submitted by horse owners were shipped overnight on ice and centrifuged (1000 x g for 10 minutes) the following morning on arrival. Plasma and serum were stored at -80°C. Samples collected by researchers were processed immediately following sample collection.

2.3.3 EMS trait measurements

Measures of obesity

Regional and generalized adiposity were measured using body condition score (BCS), girth to height ratio and neck circumference to height ratio, respectively. Body condition scores were assessed using the system developed by Henneke.[118] Body condition scores were assessed by a single investigator (NS) for researcher collected samples and scored by the same investigator from a standardized set of digital photos for samples submitted by horse owners. Horses with a body condition score ≥ 7 were considered obese.[118] A standard set of morphometric measurements were collected for each horse. The neck circumference was measured at one-half the distance between the poll and the withers with the neck held in a relaxed position.[98] To minimize variations due to muscling and hind leg conformational differences, body length was measured as the length from the point of shoulder (intermediate tubercle of the humerus) to the point of the buttock (ischiatic tuberosity), by visualizing a perpendicular line that was drawn at the point of the buttock.[119] Height was measured as the distance from the floor to the height at the third thoracic vertebra. Heart girth circumference was measured at the third thoracic vertebra. Neck circumference to height ratios (NH) were calculated by dividing the neck circumference measured halfway between the poll and withers by the withers height. Girth to height ratios (GH) were calculated by dividing the girth

(measured at the heart) by the withers height. Bodyweight was estimated using: Weight (kg) = girth (cm)2 x length (cm)/11,877.[120]

Glucose assay

Glucose (GLU) concentrations were determined using YSI glucose and lactate analyzer (2300 STAT Plus): An enzyme specific for glucose oxidase is immobilized between two membrane layers, polycarbonate and cellulose acetate. The substrate is oxidized as it enters the enzyme layer, producing hydrogen peroxide, which passes through cellulose acetate to a platinum electrode, where the hydrogen peroxide is oxidized. The resulting current is proportional to the concentration of the substrate.

Insulin assay

Insulin (INS) concentrations were determined using Coat-A-Count from Siemens per manufacturer's instructions and validated for equine serum samples as previously validated by Borer-Weir et al, 2012.[121]

Triglycerides assay

Serum triglyceride (TG) concentrations were determined using Serum Triglyceride Determination kit (TR0100) from Sigma. Assays were run per manufacturer's protocol but modified for use in a 96-well microplate format. Briefly, 200 µl of Free Glycerol Reagent was added to 2.5 µl of standards (serial dilution of Glycerol Standard Solution, G7793, Sigma), and to samples and a control serum sample and incubated for 5 min at 37C. Absorbance 1 (Abs1) was read at 540 nm using a SpectraMax 340PC microplate reader. This measurement served as the standard/sample blank. Next, 50 µl of Triglyceride Reagent was added to each well and incubated at 37°C for an additional 5 min. Following the incubation, absorbance 2 (Abs2) was read at 540 nm. The final absorbance for each well was determined by subtracting Abs1 from Abs2. Triglycerides concentration of each sample was then determined using the constructed linear standard curve (absorbance vs. concentration). Standards and samples were assayed in

triplicates.

Nonesterified fatty acids assay

Non-esterified fatty acid (NEFA) concentrations were determined using NEFA-HR from Wako. Nonesterified fatty acid concentrations were quantified by the enzymatic colorimetric method, ACS-ACOD (NEFA-HR), as outlined by Panzani et al, 2012,[122] and modified for use in a 96-well microplate format per manufacturer's instructions. Briefly, 200 µl of color reagent A solution was added to 5.0 µl of standards (serial dilution of NEFA Standard solution 276-76491), and to samples and a control serum sample and incubated for 10 min at 37°C. Absorbance 1 (Abs1) was read at 550/660 nm using a SpectraMax 340PC microplate reader. This measurement served as the standard/sample blank. Next, 100 µl of color reagent B solution was added to each well and incubated at 37°C for an additional 10 min. Following the incubation, absorbance 2 (Abs2) was read at 550/660 nm. The final absorbance for each well was determined by subtracting Abs1 from Abs2. Nonesterified fatty acid concentration of each sample was then determined using the constructed linear standard curve (absorbance vs. concentration). Standards and samples were assayed in duplicate.

Adrenocorticotropin hormone assay

Adrenocorticotropin hormone (ACTH) concentrations were determined from plasma samples using an automated chemiluminescent enzyme immunoassay system (Immulite) validated previously by Perkins et al.[123] on the Immulite 1000 (Siemens). Intra assay variation at the low control (29 pg/ml) was 3% and 2.5% at the high control (367 pg/ml) for 10 replications.

Leptin assay

Leptin (LEP) concentrations were determined using Multi-Species Leptin Radioimmunoassay (RIA) (XL-85K) from EMD Millipore per manufacturer's instructions. This assay was validated previously by Fitzgerald, B.P., McManus, C.J., (2000),[124] and McManus and Fitzgerald (2003).[125] The kit utilized 125I-labeled recombinant human leptin, a guinea pig multispecies leptin primary antibody, and a goat anti-guinea pig IgG

serum for the precipitating reagent. Purified recombinant human leptin was used for the kit standards and quality controls. Samples were run in duplicate and counted for one minute in a gamma counter.

Adiponectin assay

Serum (APN) adiponectin concentrations were determined using a human high molecular weight (HMW) adiponectin enzyme-linked immunosorbent assay (ELISA) (EZHMWA-64K) from EMD Millipore following a protocol validated for equine serum by Wooldridge et al 2012.[101] Samples were run in duplicate.

2.3.4 Signalment and medical history

Signalment information including age, breed, sex, and pregnancy status was documented. Clinical group status (non-obese/no prior laminitis history, non-obese/prior laminitis history, obese/no prior laminitis history, and obese/prior laminitis history) was determined based on medical history and BCS. Supplementation with L-thyroxine was also documented.

2.3.5 Environment measurements

Geographic latitudes (determined from address information[a]) and month of sample collection were documented. Exercise data included average hours of stall confinement per day, average hours of exercise per week, and description of the type of exercise (see Table A.2). Average hours per day grazing was also reported. Hay, pasture, and grain samples were obtained for dietary analysis. Hay samples were collected from 5 bales of similar hay (same species and cutting) by opening each bale and grabbing a fist full of hay. These samples were mixed and then a subsample (1 quart sized bag) collected for testing. The total amount of hay consumed per horse each day was determined by multiplying the number of flakes of hay consumed per day by the weight per flake of hay. Weight per flake of hay was determined by weighing of 10 flakes and determining an average weight.

Pasture samples were collected from pastures grazed by horses by walking a 'Z' or

[a]http://www.findlatitudeandlongitude.com/

'M' pattern throughout the entire pasture and collection of 10 random samples. Samples included all forage in a one-foot square area cut down to the ground. Areas where horses had defecated or areas where no grazing occurred were avoided. These samples were then placed into a bucket and mixed thoroughly and a subsample (1 quart sized bag) was collected for testing. Because standing forage is approximately 85% water, the samples were kept out of direct sunlight and placed in a cooled box for shipping to the laboratory. All diet samples were stored at -20° or -80°C until being shipped to Equi-Analytical Laboratories for analysis (pasture samples were shipped to Equi-Analytical Laboratories on dry ice).

Daily caloric consumption (Mcal) and grams crude protein (CP), neutral detergent fiber (NDF), starch, and water soluble carbohydrates (WSC) consumed per kg bodyweight (bwt) were calculated based on dry matter weight of each dietary component determined from analysis of dietary samples multiplied by the daily dry matter (DM) weight fed of each dietary component. Dry matter weight of daily pasture consumption was calculated by multiplying the determined pasture dry matter proportion by an estimate of the amount of pasture consumed daily based on time spent on pasture. Grams of dry matter intake per hour spent on pasture was extrapolated from data indicating horses on pasture for 3,6,9, and 24 hours consumed 1.96, 1.52, 1.12, and 0.57 g DM intake per kg bwt per hour.[126] A linear regression model was fit with log transformed DM intake values as an outcome and hours on pasture as a predictor to determine the following prediction estimate equation (r^2=0.95).

$$\text{g dry matter intake per hour} = (10^{[(\text{hrs grazing}*-0.0244)+0.3265]})* \text{ hrs grazing} * \text{kg bwt} \tag{2.1}$$

Gram dry matter intake estimates where reduced by 75% for horses reported to have worn grazing muzzles.[127]

2.3.6 Statistical Methods

A multilevel, multivariate response model in which multiple outcomes are simultaneously regressed over a set of explanatory variables was constructed to answer the research questions.[128] See Section 1.3 for an overview of multivariate, multilevel models.

A multilevel linear model (aka hierarchical linear model) is a standard method for analyzing outcome data with hierarchical structure such as observations clustered within individuals and individuals clustered within groups. The hierarchical structure leads to correlations among the observations within a cluster and therefore violation of assumptions of outcome independence leading to inaccurate standard errors and significance test of fixed effects. Multilevel models account for the correlation among observations by including additional error terms (random effects).

By including multiple outcomes within the multivariate response model, it is possible to estimate the covariance between different outcomes nested within individuals and within farms, as well as the variance for each outcome at both the individual and farm level in a simultaneous manner, see Equation 1.13. Additionally, the Akaike Information Criterion (AIC)[114] and the likelihood ratio test indicated a multivariate response model provided a better fit than a univariate response model.

First, a "Null" model was fit that contained no explanatory variables in order to estimate the amount of metabolic trait variation and co-variation present at the farm and individual horse level. The multivariate response variables included neck circumference to height ratio (NH), girth to height ratio (GH), fasting glucose (GLU), fasting insulin (INS), glucose 75 minutes post oral sugar challenge (GLU OST), insulin 75 minutes post oral sugar challenge (INS OST), fasting triglycerides (TG), fasting non-esterified fatty acids (NEFA), fasting adrenocorticotropin hormone (ACTH), fasting leptin (LEP), and fasting adiponectin (APN). INS, TG, NEFA, ACTH, LEP, and APN exhibited a skewed distribution and required log or square root transformations to achieve a more normal distribution with a centrally located mean.

Secondly, a "full" model containing all of the explanatory models was fit to examine the degree of reduction in variance estimates achieved with inclusion of explanatory variables, and to estimate the effect size of individual and farm related factors on trait outcomes. The full model explanatory variables included: age, breed, sex, clinical group, sampling time of year, latitude, Mcal/kg bwt per day, CP/kg bwt per day, NDF/kg bwt per day, Starch/kg bwt per day, WSC/kg bwt per day, hours of exercise per day, hours grazing per day, hours stalled per day, owner submitted versus researcher collected sample, and L-thyroxine supplementation(yes/no). Correlation estimates and standard errors are directly provided by the ASReml-R software[129] used to perform the analyses,

95% confidence intervals were calculated as the estimate ±1.96 x standard error. Fixed effects are expressed as standardized coefficients (represents the effect size in terms of the number of standard deviations from the reference group, i.e. the grand mean) and their standard error. The Wald z-test was used to test the significance of a single fixed effect and a Wald chi-square test was utilized to test the overall significance level of a categorical explanatory variable such as month of year. The percent reduction in the "Null" model variance estimate with inclusion of explanatory variables can be thought of as a pseudo-R squared and an indicator of the model "goodness of fit". This was used to determine the proportion of variation explained by explanatory predictor variables included in the model.

The fractional importance of individual predictor variables in explaining metabolic trait measures can be estimated by calculating the percent reduction in the variance estimate for a model that contained all predictor variables relative to a model that eliminated the predictor variable(s) of interest. However, this approach will not yield individual explanatory variable estimates of variance explained that sum up to the total amount of explained variance by the "full" model if correlation exists among the explanatory variables. In order to obtain estimates of variance explained by each individual explanatory variable that sum up to the total amount of variance explained by the "full" model, a Shapley value regression approach was utilized.[130] Shapley value regression involves fitting numerous models with each possible combination of the explanatory predictor variables ranging from a single explanatory predictor included in the model to inclusion of all explanatory variables. The Shapley Value is calculated across all possible combinations of predictors and can be defined as:

$$ShapleyValue_j = \sum_k \sum_i \gamma_k [\nu(M_{i|j}) - \nu(M_{i|j(-j)})]$$

Where:

$\nu(M_{i|j})$ is the pseudo-R^2 of a model i containing predictor j and

$\nu(M_{i|j(-j)})$ is the pseudo-R^2 of the same model i without predictor j. (2.2)

$\gamma_k = \dfrac{k!(n-k-1)!}{n!}$ is a weight based on the number of predictors in total (n) and the number of predictors in this model (k).

For the Shapley Value analysis, the explanatory variable clinical group was fit as two separate explanatory variables 1) obesity status and 2) prior laminitis status in order to obtain separate estimates of the variance explained obesity and the variance explained by laminitis versus only obtaining an estimate of the variance explained by clinical group.

2.4 Results

2.4.1 Study population descriptive statistics and demographics

Descriptive summary statistics of the metabolic trait response variables NH, GH, GLU, INS, GLU OST, INS OST, TG, NEFA, ACTH, LEP, and APN are presented in Table 2.1. Descriptive summary statistics for metabolic trait responses for each level of a categorical explanatory variable are presented in Appendix table **??**. INS, TG, NEFA, ACTH, LEP, and APN exhibited a skewed distribution and required log or square root transformations to achieve a more normal distribution with a centrally located mean (mean approximates median). Summary statistics describing continuous explanatory variables for the population are presented in Table 2.2 and a summary of categorical explanatory variable population demographics are presented in Table 2.3.

Table 2.1: Descriptive summary statistics for metabolic trait outcome variables measured in the study population

Trait	N	Mean	SD	Median	Range
NH	603	0.66	0.05	0.66	0.50-0.83
GH	608	1.22	0.07	1.22	0.77-1.49
GLU (mg/dl)	608	77.0	10.4	76.8	32.8-153.0
INS (μIU/ml)	609	12.5	35.1	6.7	1.5-632.0
logINS (μIU/ml)	609	0.84	0.41	0.83	0.18-2.80
GLU OST (mg/dl)	513	98.0	18.4	97.4	40.7-203.0
INS OST (μIU/ml)	514	40.0	67.5	22.3	1.5-783.6
logINS OST (μIU/ml)	514	1.35	0.45	1.35	0.18-2.89
TG (mg/dl)	602	31.7	26.8	25.4	3.7-337.3
logTG (mg/dl)	602	1.42	0.26	1.40	0.57-2.53
NEFA (mmol/L)	607	0.23	0.18	0.19	0.00-1.22
$\surd$NEFA (mmol/L)	607	0.45	0.19	0.44	0.00-1.10
ACTH (pg/ml)	608	35.4	32.5	26.8	10.0-287.0
logACTH (pg/ml)	608	1.47	0.24	1.43	1.00-2.46
LEP (ng/ml)	600	6.03	3.91	5.18	0.00-26.15
$\surd$LEP (ng/ml)	600	2.33	0.78	2.28	0.00-5.11
APN (ng/ml)	609	4349	3088	3728	0-22189
$\surd$APN (ng/ml)	609	61.5	23.7	61.1	0.0-149.0

NH: neck circumference to height ratio, GH: girth to height ratio, GLU: glucose, INS: insulin, log: base 10 log transformation, OST: oral sugar test, sample was obtained 75 minutes following oral sugar administration, $\surd$: square root transformation, ACTH: adrenocorticotropin hormone, LEP: leptin, APN: adiponectin

Table 2.2: Descriptive summary statistics for continuous explanatory variables measured in the study population

	N	Mean	SD	Median	Range
WSC g per kg bwt	610	2.0	0.9	1.8	0.2-5.5
Starch g per kg bwt	610	0.6	0.6	0.4	0.0-5.2
NDF g per kg bwt	610	11.2	5.0	9.8	1.7-27.7
CP g per kg bwt	610	2.7	1.3	2.4	0.5-7.3
Mcal g per kg bwt	610	0.042	0.016	0.039	0.007-0.104
Hrs grazing/day	610	10.7	10.7	8.0	0.0-24.0
Hrs exercise/week	610	1.4	2.4	0.0	0.0-20.0
Hrs stall/day	610	3.5	6.7	0.0	0.0-24.0
Latitude	610	40.776	5.022	41.821	26.715-51.314
Age	610	12.6	6.1	12.0	2.0-33.0

WSC: water soluble carbohydrate, bwt: bodyweight, NDF: neutral detergent fiber, CP: crude protein, Mcal: megacalorie, Hrs: hours

Table 2.3: Demographic summary of categorical explanatory variables measured in the study population

Parameter	Categorical Variable	N
Non-obese, no prior laminitis	Group	340
Non-obese, prior laminitis	Group	95
Obese, no prior laminitis	Group	111
Obese, prior laminitis	Group	64
Morgan	Breed	288
Arab	Breed	63
Pony	Breed	96
Quarter Horse	Breed	57
Tennessee Walking Horse	Breed	45
HR (other high risk breeds)	Breed	45
LR (other low risk breeds)	Breed	16
Gelding	Sex	210
Mare	Sex	368
Stallion	Sex	32
Owner submitted sample	Sample handling	277
Researcher collected sample	Sample handling	333
No L-thyroxine supplementation	L-thyroxine supplementation	586
Yes L-thyroxine supplementation	L-thyroxine supplementation	24
DEC collection	Collection month	58
JAN collection	Collection month	8
FEB collection	Collection month	39
MAR collection	Collection month	86
APR collection	Collection month	69
MAY collection	Collection month	73
JUN collection	Collection month	11
JUL collection	Collection month	38
AUG collection	Collection month	100
SEP collection	Collection month	60
OCT collection	Collection month	52
NOV collection	Collection month	16

Horses with a body condition score <7 were classified as non-obese, horses with a body condition score ≥ 7 were classified as obese. High risk breeds includes Mustang, Warmblood, Haflinger, Fox Trotter, Rocky Mountain Horse, Saddlebred, Paso Fino, and Peruvian Paso. Low risk breeds includes Thoroughbred and draft breeds. High risk and low risk breeds were subjectively determined based on clinical judgement/anecdotal prevalence reports. The Pony breed group is 88.5%Welsh, 5.2%Welsh cross, and 6.3% other pony breeds.

2.4.2 Variance decomposition of equine metabolic traits at the farm and individual horse level

Estimates of metabolic trait variance at the farm and individual horse level from the Null model (no explanatory variables included) are presented in Table 2.4. The 0.23-0.49 proportion of variation present at the farm level for the metabolic traits indicate homogeneity of traits among horses within a farm or intracluster correlation, i.e. metabolic traits of two horses selected from the same farm are likely to be more similar than metabolic traits of two horses selected from different farms. Presence of intracluster correlation indicate farm related characteristics shared by horses sampled from the same farm such as diet or season at time of sampling may contribute to variation of metabolic traits. For all traits, the majority of the variation (0.51-0.77) is present at the individual level indicating that the metabolic measures of individual horses within a farm vary about the farm mean due to individual level characteristics, e.g. age or sex.

Table 2.4: Individidual and farm level variance estimates (standard error) for metabolic phenotypes

	individual level variance	farm level variance
NH	0.57(0.04)	0.5(0.08)
	53%	**47%**
GH	0.6(0.04)	0.45(0.08)
	57%	**43%**
GLU	0.69(0.05)	0.37(0.07)
	65%	**35%**
logINS	0.69(0.05)	0.39(0.08)
	64%	**36%**
GLU OST	0.61(0.04)	0.45(0.09)
	58%	**42%**
logINS OST	0.71(0.05)	0.35(0.07)
	67%	**33%**
logTG	0.71(0.05)	0.22(0.05)
	76%	**24%**
$\sqrt{}$NEFA	0.56(0.04)	0.45(0.08)
	55%	**45%**
ACTH	0.56(0.04)	0.54(0.09)
	51%	**49%**
$\sqrt{}$LEP	0.71(0.05)	0.33(0.07)
	68%	**32%**
$\sqrt{}$APN	0.78(0.05)	0.23(0.05)
	77%	**23%**

Phenotypes were scaled to zero mean and variance equal to one standard deviation.

2.4.3 Correlation of equine metabolic traits at the farm and individual horse level

The correlation (trait co-variation) between the 11 measured outcome variables (i.e. NH, GH, GLU, INS, GLU OST, INS OST, TG, NEFA, ACTH, LEP, and APN) was incorporated in the model at both the individual and farm level. Estimates of metabolic trait correlation from the "Null" model at the individual horse level are presented in Table 2.5 and estimates of metabolic trait correlation from the "Null" model at the farm level are presented in Table 2.6 (covariance estimates are reported in Appendix Tables A.3 to A.4). Cohen's guidelines[131]were used to assess the correlation coefficient strength (>0.5=strong, 0.3-0.5= moderate, 0.1-0.3=weak).

Trait correlation at the individual level (see Table 2.5) suggests that two measures are physiologically related due to physiologic factors or characteristics of the individual (e.g. age, sex). The absolute values of trait correlation coefficients at the individual level ranged from zero correlation to 0.67. An individual horse's post-OST GLU or INS levels were strongly correlated with the individual horse's fasted levels. Fasted GLU exhibited moderate, strong correlation with fasted INS, as did post-OST GLU with post-OST INS. An individual horse's TG was moderately correlated with the horse's INS (fasted and post-OST) level. An individual's LEP level exhibited a moderate, strong positive correlation with INS (fasted and post-OST) where as an individual's APN level exhibited a moderate, strong negative correlation with INS (fasted and post-OST) and TG. NH, a measure of regional adiposity was moderately correlated with INS and demonstrated a weak negative correlation with APN. GH, a measure of regional adiposity was strongly correlated with LEP. APN and LPN exhibited a very weak correlation.

Correlations between traits at the farm level (see Table 2.6) suggest that shared farm level factors, such as environment, season of sampling, etc are impacting both traits. The absolute values of trait correlation coefficients at the farm level ranged from zero correlation to 0.86. At the farm level, post-oral sugar test GLU or INS levels were very strongly correlated with the farm's mean fasted levels. Farm mean post-OST GLU were very strongly correlated with the farm's mean post-OST INS levels where as farm fasted GLU levels were moderately correlated with farm's fasted INS levels. Farm TG levels were strongly correlated with farm's fasting INS and moderately correlated with the farm's post OST INS. Farm ACTH levels exhibited a strong positive

correlation with the farm's post-OST INS and a strong negative correlation with APN. Weak moderate, positive correlation existed for farm ACTH levels and farm fasting INS, TG, post-OST GLU. Farm LEP levels were strongly correlated with farm INS (fasted and post-OST) levels. Farm APN demonstrated strong negative correlations with GLU (fasted and post-OST), INS (fasted and post-OST), and TG, which were stronger than the correlations at the individual level. In contrast, APN showed a moderate positive correlation with NH, which is the opposite of the moderate negative correlation between APN and NH at the individual level. Some traits exhibit correlation on one level but not the other, for example APN is correlated at the farm level with ACTH but not at the individual level.

Table 2.5: Individual level correlation estimates, 95% confidence intervals, and p-values for metabolic phenotypes determined from the "null" model.

	NH	GH	GLU	INS	GLU OST	INS OST	TG	NEFA	ACTH	LEP
GH	**0.40** **0.32,0.48** **3.0e-23**									
GLU	-0.01 -0.10,0.09 8.8e-01	0.00 -0.09,0.09 9.8e-01								
INS	**0.30** **0.21,0.38** **7.0e-12**	**0.22** **0.13,0.31** **1.3e-06**	**0.38** **0.30,0.46** **8.0e-21**							
GLU OST	0.03 -0.07,0.13 5.6e-01	0.04 -0.06,0.13 4.7e-01	**0.51** **0.44,0.59** **5.6e-45**	**0.24** **0.15,0.33** **1.6e-07**						
INS OST	**0.32** **0.23,0.40** **5.7e-13**	**0.22** **0.13,0.31** **2.2e-06**	**0.29** **0.20,0.37** **1.5e-10**	**0.67** **0.62,0.73** **1.5e-141**	**0.42** **0.34,0.50** **1.5e-25**					
TG	**0.21** **0.13,0.30** **1.8e-06**	**0.11** **0.01,0.20** **2.3e-02**	**0.12** **0.03,0.21** **1.2e-02**	**0.33** **0.25,0.41** **1.8e-15**	**0.13** **0.03,0.22** **9.0e-03**	**0.38** **0.30,0.46** **9.3e-20**				
NEFA	-0.01 -0.10,0.08 8.2e-01	-0.02 -0.11,0.08 7.2e-01	**-0.14** **-0.23,-0.05** **3.2e-03**	**-0.09** **-0.19,0.00** **4.4e-02**	**0.10** **0.00,0.19** **4.0e-02**	-0.06 -0.16,0.03 1.8e-01	-0.03 -0.12,0.06 5.5e-01			
ACTH	0.06 -0.03,0.16 1.9e-01	-0.02 -0.11,0.07 6.4e-01	0.05 -0.05,0.14 3.3e-01	**0.15** **0.06,0.24** **1.1e-03**	0.04 -0.05,0.14 3.7e-01	0.08 -0.02,0.17 1.1e-01	0.07 -0.03,0.16 1.6e-01	0.09 0.00,0.18 5.0e-02		
LEP	**0.28** **0.20,0.37** **9.7e-11**	**0.43** **0.36,0.51** **3.4e-29**	**0.16** **0.07,0.25** **3.5e-04**	**0.34** **0.26,0.42** **1.9e-16**	**0.20** **0.10,0.29** **2.8e-05**	**0.40** **0.32,0.48** **2.2e-23**	**0.23** **0.14,0.32** **2.1e-07**	**-0.18** **-0.27,-0.09** **8.2e-05**	-0.01 -0.10,0.08 8.2e-01	
APN	**-0.21** **-0.30,-0.12** **2.5e-06**	-0.08 -0.17,0.01 7.6e-02	**-0.09** **-0.18,0.00** **4.5e-02**	**-0.35** **-0.43,-0.27** **3.1e-18**	-0.07 -0.17,0.02 1.3e-01	**-0.37** **-0.45,-0.28** **1.3e-18**	**-0.41** **-0.49,-0.34** **1.4e-27**	**0.12** **0.03,0.21** **1.0e-02**	0.00 -0.09,0.09 9.7e-01	**-0.11** **-0.20,-0.03** **1.2e-02**

Correlation
(-1.0,-0.6)
(-0.6,-0.4)
(-0.4,-0.2)
(-0.2,0.0)
(0.0,0.2)
(0.2,0.4)
(0.4,0.6)
(0.6,1.0)

Bold face cells indicate p-value<0.05. INS, INS OST, TG, and ACTH were log transformed. NEFA, LEP, and APN were square root transformed.

Table 2.6: Farm level correlation estimates, 95% confidence intervals, and p-values for metabolic phenotypes determined from the "null" model.

	NH	GH	GLU	INS	GLU OST	INS OST	TG	NEFA	ACTH	LEP
GH	**0.50** **0.31,0.68** **1.3e-07**									
GLU	-0.03 -0.28,0.22 8.0e-01	-0.09 -0.35,0.16 4.8e-01								
INS	0.00 -0.26,0.25 9.7e-01	-0.01 -0.27,0.25 9.5e-01	**0.31** **0.06,0.55** **1.5e-02**							
GLU OST	0.11 -0.14,0.36 3.9e-01	0.14 -0.12,0.39 2.9e-01	**0.78** **0.65,0.90** **5.2e-33**	**0.44** **0.21,0.67** **1.8e-04**						
INS OST	-0.03 -0.30,0.24 8.1e-01	0.07 -0.20,0.34 6.1e-01	**0.40** **0.15,0.64** **1.3e-03**	**0.86** **0.76,0.96** **3.5e-68**	**0.70** **0.54,0.86** **3.9e-18**					
TG	0.05 -0.23,0.33 7.3e-01	0.02 -0.26,0.31 8.7e-01	0.14 -0.15,0.43 3.4e-01	**0.44** **0.19,0.69** **5.7e-04**	0.25 -0.03,0.53 8.5e-02	**0.30** **0.02,0.58** **3.3e-02**				
NEFA	0.08 -0.15,0.31 5.1e-01	**0.27** **0.05,0.50** **1.9e-02**	0.03 -0.22,0.28 8.1e-01	-0.07 -0.32,0.18 5.9e-01	**0.38** **0.15,0.60** **1.3e-03**	0.23 -0.02,0.48 7.4e-02	0.08 -0.20,0.36 5.7e-01			
ACTH	-0.17 -0.40,0.06 1.4e-01	-0.11 -0.34,0.12 3.4e-01	**0.26** **0.02,0.50** **3.2e-02**	**0.36** **0.13,0.58** **1.8e-03**	**0.35** **0.12,0.58** **3.0e-03**	**0.52** **0.30,0.74** **2.3e-06**	**0.32** **0.06,0.58** **1.7e-02**	**0.26** **0.04,0.48** **1.8e-02**		
LEP	0.13 -0.12,0.38 3.2e-01	0.07 -0.19,0.33 5.9e-01	-0.23 -0.50,0.05 1.1e-01	**0.48** **0.26,0.70** **2.4e-05**	0.05 -0.22,0.33 7.0e-01	**0.43** **0.19,0.67** **4.1e-04**	0.14 -0.16,0.43 3.5e-01	0.10 -0.16,0.36 4.5e-01	0.07 -0.19,0.32 6.0e-01	
APN	**0.39** **0.13,0.66** **3.6e-03**	0.27 0.00,0.54 5.1e-02	**-0.56** **-0.79,-0.33** **2.3e-06**	**-0.53** **-0.76,-0.30** **5.1e-06**	**-0.45** **-0.70,-0.19** **5.5e-04**	**-0.50** **-0.75,-0.25** **7.5e-05**	**-0.47** **-0.72,-0.22** **2.5e-04**	-0.03 -0.31,0.25 8.3e-01	**-0.46** **-0.71,-0.20** **3.9e-04**	-0.09 -0.39,0.21 5.4e-01

Correlation
(-1.0,-0.6)
(-0.6,-0.4)
(-0.4,-0.2)
(-0.2,0.0)
(0.0,0.2)
(0.2,0.4)
(0.4,0.6)
(0.6,1.0)

Bold face cells indicate p-value<0.05. INS, INS OST, TG, and ACTH were log transformed. NEFA, LEP, and APN were square root transformed.

2.4.4 Effect of physiologic factors (age, sex, breed, and clinical group status) on metabolic trait variation

The significance values from a Wald chi-square test of the overall effect of age, sex, breed, and clinical group status on metabolic trait variation from the full model are presented in Table 2.7. Fixed effect estimates for the explanatory variables age, sex, breed, and clinical group status from the full model are reported in Appendix tables A.5-A.15 and illustrated in Figure 2.1. Additional covariates/explanatory variables in the full model included diet composition (Mcal, CP, NDF, Starch, and WSC per kg bwt per day), average hours per day grazing, average hours per day exercising, average hours per day stalled, month at time of sampling, farm latitude, thyroxine supplementation (yes/no), and researcher collected sample vs owner submitted. Figure 2.1 illustrates the effect estimates of age, sex, breed, and clinical group in terms of standard deviations from the reference group indicated by the red vertical line (non-obese, 12 year old Morgan geldings with no prior history of laminitis). Several significant effects were identified for numerous traits and are described in more detail below.

The predicted metabolic trait means for mares, geldings, and stallions determined from the full model are reported in Table 2.8. Significant overall and pairwise effects of sex on metabolic trait variation included NH (stallion > gelding > mare), GH (geldings, mares > stallion), GLU fasted and OST (stallions > mares), OST INS (stallion > gelding > mare), and LEP (mares > geldings).

The predicted metabolic trait means for the different breeds are reported in Table 2.9. Significant overall and pairwise effects of breed on metabolic trait variation included NH (ponies, Morgan > low risk breeds, Morgan > Arabian), GH (QH > low risk breeds), INS (Morgan, TW, HR > QH and Ponies > QH, Arabs) and OST INS (Morgan, TW, HR, ponies > QH), ACTH (ponies > QH, TW, Morgan), LEP (Morgan, ponies, TW, high risk breeds > QH, Morgan > low risk breeds), and APN (QH > Arabians).

The predicted metabolic trait means according to obesity and/or prior laminitis status are reported in Table 2.10. Significant overall and pairwise effects of obesity and/or prior laminitis status on metabolic trait variation included NH ($\text{obese}^{+}\text{laminitis}^{+}$ > $\text{obese}^{+}\text{laminitis}^{-}$ > $\text{obese}^{-}\text{laminitis}^{+}$ > $\text{obese}^{-}\text{laminitis}^{-}$), GH ($\text{obese}^{+}\text{laminitis}^{+}$, $\text{obese}^{+}\text{laminitis}^{-}$ > $\text{obese}^{-}\text{laminitis}^{+}$, $\text{obese}^{-}\text{laminitis}^{-}$), INS ($\text{obese}^{+}\text{laminitis}^{+}$ > $\text{obese}^{+}\text{laminitis}^{-}$, $\text{obese}^{-}\text{laminitis}^{+}$ > $\text{obese}^{-}\text{laminitis}^{-}$) and OST INS

(obese$^+$laminitis$^+$, obese$^+$laminitis$^-$, obese$^-$laminitis$^+$ > obese$^-$laminitis$^-$), TG (obese$^+$laminitis$^+$, obese$^-$laminitis$^+$ > obese$^+$laminitis$^-$ > obese$^-$laminitis$^-$), LEP (obese$^+$laminitis$^+$, obese$^+$laminitis$^-$ > obese$^-$laminitis$^+$, obese$^-$laminitis$^-$), and APN (obese$^-$laminitis$^-$, obese$^+$laminitis$^-$ > obese$^-$laminitis$^+$, obese$^+$laminitis$^+$).

The predicted effect of age on the median trajectory of insulin and ACTH levels is illustrated in Figure 2.2. Age correlated positively with INS (fasted and OST) and ACTH levels.

Table 2.7: Significance results for overall association of physiologic factors (age, breed, sex, and clinical group) with metabolic trait variation determined from the full multivariate, multilevel model

	+ 1 SD age	sex	breed	clinical group
NH	7.0e-01	**2.1e-07**	**7.6e-04**	**9.0e-22**
GH	5.2e-02	**1.2e-04**	**9.2e-03**	**5.0e-18**
GLU	3.6e-01	**2.7e-03**	1.8e-01	5.8e-01
logINS	**7.4e-04**	2.2e-01	**1.4e-05**	**2.3e-17**
GLU OST	6.1e-01	**1.2e-02**	6.0e-01	3.2e-01
logINS OST	**1.1e-02**	**3.9e-04**	**2.1e-04**	**8.2e-12**
logTG	5.4e-02	6.5e-02	**2.7e-02**	**7.1e-09**
$\sqrt{}$NEFA	9.2e-01	5.4e-01	4.7e-01	9.1e-01
logACTH	**9.6e-20**	3.6e-01	**7.0e-05**	4.4e-01
$\sqrt{}$Leptin	9.9e-01	**3.9e-03**	**5.2e-08**	**5.9e-11**
$\sqrt{}$Adiponectin	3.6e-01	4.6e-01	**4.7e-02**	**6.6e-10**

Figure 2.1: Estimated effect of age, sex, breed, and clinical group status on metabolic trait variation. Vertical line at zero standard deviations indicates the reference group[non-obese Morgan geldings with no history of laminitis and age equal to the mean population age].

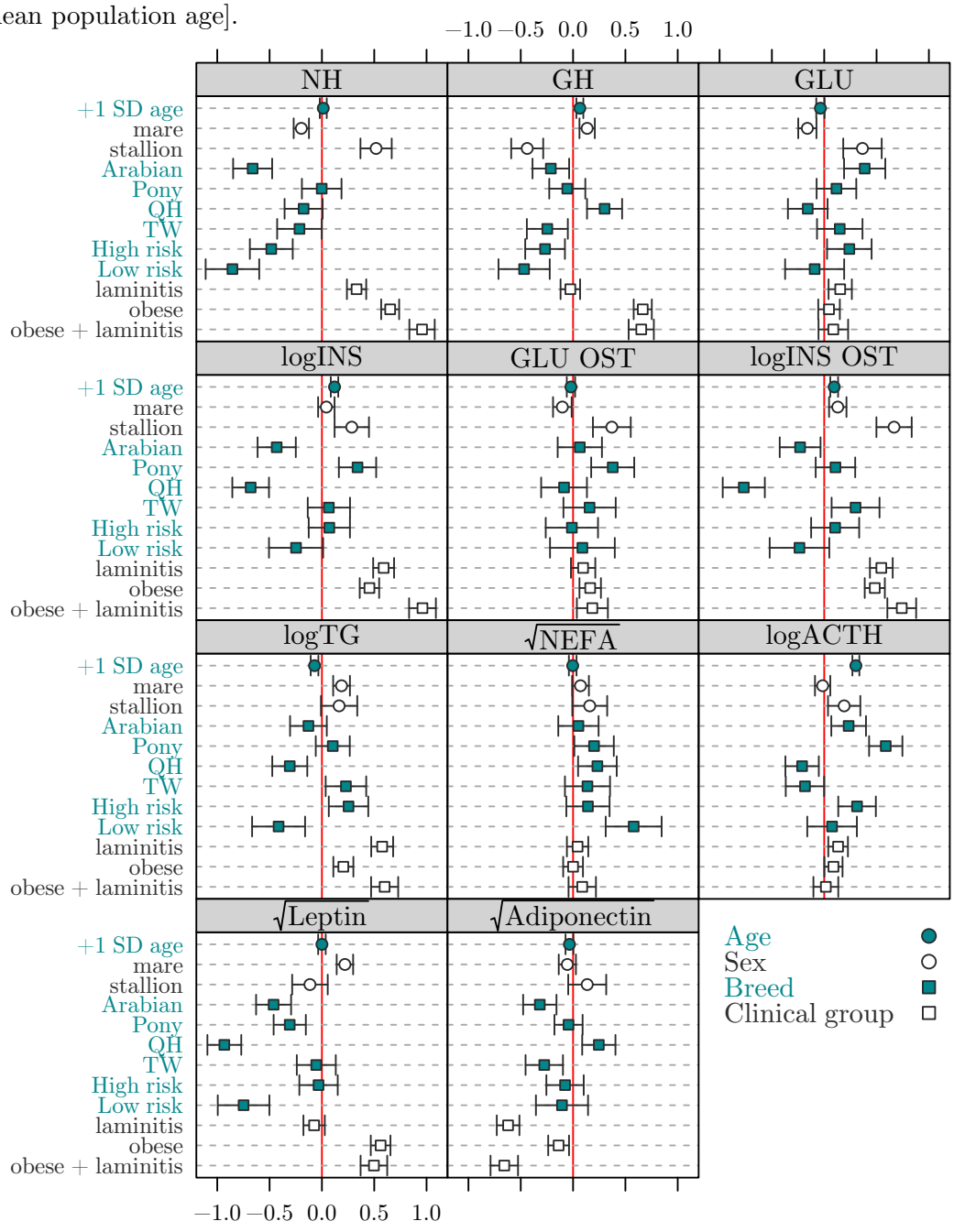


Table 2.8: Predicted metabolic trait central means [95% confidence interval] by sex determined from the full multivariate, multilevel model

NH		**GH**		**GLU (mg/dl)**		**INS (μIU/ml)**	
mare	0.65[0.64,0.67]a	stallion	1.19[1.16,1.21]a	mare	76.1[73.3,78.9]a	gelding	7.9[6.2,10.1]a
gelding	0.66[0.65,0.68]b	gelding	1.22[1.20,1.23]b	gelding	77.8[74.9,80.7]ab	mare	8.2[6.5,10.4]a
stallion	0.69[0.67,0.71]c	mare	1.23[1.21,1.24]b	stallion	81.6[77.3,85.9]b	stallion	10.4[7.3,14.8]a
GLU OST (mg/dl)		**INS OST (μIU/ml)**		**TG (mg/dl)**		**NEFA (mmol/L)**	
mare	97.7[92.2,103.2]a	gelding	24.5[18.2,32.8]a	gelding	27.7[24.0,32.0]a	gelding	0.20[0.16,0.25]a
gelding	99.6[94.0,105.2]ab	mare	28.0[21.0,37.2]a	stallion	30.6[24.4,38.3]a	mare	0.21[0.17,0.26]a
stallion	106.4[98.5,114.3]b	stallion	49.1[32.5,74.4]b	mare	31.0[27.0,35.6]a	stallion	0.23[0.16,0.30]a
ACTH (pg/ml)		**LEP (ng/ml)**		**APN (ng/ml)**			
mare	27.4[24.2,30.9]a	stallion	4.54[3.39,5.87]ab	mare	2852[2312,3448]a		
gelding	27.6[24.3,31.3]a	gelding	4.93[4.14,5.79]a	gelding	2992[2422,3622]a		
stallion	30.6[25.3,37.0]a	mare	5.72[4.89,6.61]b	stallion	3353[2374,4501]a		

Different letters within a column indicate a statistically significant pairwise difference ($p<0.05$ with holm adjustment for multiple comparison).

Table 2.9: Predicted metabolic trait central means [95% confidence interval] by breed determined from the full multivariate, multilevel model

NH		**GH**		**GLU (mg/dl)**		**INS (μIU/ml)**	
LR	0.64[0.61,0.67]a	LR	1.19[1.15,1.22]a	QH	75.9[71.7,80.1]a	QH	5.2[3.7,7.4]a
Arab	0.65[0.63,0.67]ab	HR	1.20[1.17,1.23]ab	LR	76.6[70.6,82.5]a	Arab	6.6[4.6,9.4]ab
HR	0.66[0.64,0.68]abc	TW	1.20[1.18,1.23]ab	Morgan	77.5[74.3,80.8]a	LR	7.8[4.8,12.8]abc
TW	0.68[0.65,0.70]abc	Arab	1.20[1.18,1.23]ab	Pony	78.8[74.7,82.8]a	Morgan	9.9[7.5,13.0]bc
QH	0.68[0.66,0.70]abc	Pony	1.22[1.19,1.24]ab	TW	79.1[74.8,83.4]a	TW	10.5[7.3,15.2]bc
Pony	0.69[0.67,0.71]bc	Morgan	1.22[1.20,1.24]ab	HR	80.0[75.9,84.2]a	HR	10.6[7.5,15.0]bc
Morgan	0.69[0.67,0.70]c	QH	1.24[1.22,1.27]b	Arab	81.6[77.4,85.8]a	Pony	13.7[9.7,19.2]c
GLU OST (mg/dl)		**INS OST (μIU/ml)**		**TG (mg/dl)**		**NEFA (mmol/L)**	
QH	98.1[89.7,106.5]a	QH	16.1[10.4,25.0]a	LR	23.8[17.5,32.2]a	Morgan	0.18[0.14,0.23]a
HR	99.5[90.9,108.0]a	LR	28.1[15.4,51.3]ab	QH	25.3[20.5,31.3]a	Arab	0.19[0.13,0.26]a
Morgan	99.7[93.4,106.0]a	Arab	28.3[18.5,43.2]ab	Arab	28.2[22.7,34.9]a	TW	0.20[0.14,0.28]a
Arab	100.9[92.8,109.0]a	Morgan	36.0[25.9,50.0]b	Morgan	30.4[25.8,35.8]a	HR	0.21[0.14,0.28]a
LR	101.3[89.8,112.8]a	HR	40.1[25.7,62.7]b	Pony	32.3[26.4,39.6]a	Pony	0.22[0.15,0.29]a
TW	102.6[93.9,111.2]a	Pony	40.2[26.7,60.6]b	TW	34.8[27.9,43.4]a	QH	0.22[0.16,0.30]a
Pony	106.6[98.8,114.5]a	TW	49.2[31.3,77.3]b	HR	35.4[28.6,43.7]a	LR	0.29[0.19,0.41]a
ACTH (pg/ml)		**LEP (ng/ml)**		**APN (ng/ml)**			
QH	23.8[19.9,28.6]a	QH	3.25[2.36,4.29]a	Arab	2474[1748,3326]a		
TW	24.2[20.0,29.3]a	LR	3.79[2.42,5.46]ab	TW	2576[1799,3492]ab		
Morgan	26.7[23.2,30.9]a	Arab	4.71[3.62,5.95]abc	LR	2999[1825,4464]ab		
LR	27.8[21.5,36.1]ab	Pony	5.25[4.15,6.47]bc	HR	3076[2243,4039]ab		
Arab	30.3[25.2,36.5]ab	TW	6.19[4.87,7.68]bc	Pony	3164[2398,4036]ab		
HR	31.7[26.4,38.1]ab	HR	6.27[4.98,7.72]bc	Morgan	3281[2609,4029]ab		
Pony	36.8[30.8,43.9]b	Morgan	6.40[5.39,7.50]c	QH	3983[3024,5073]b		

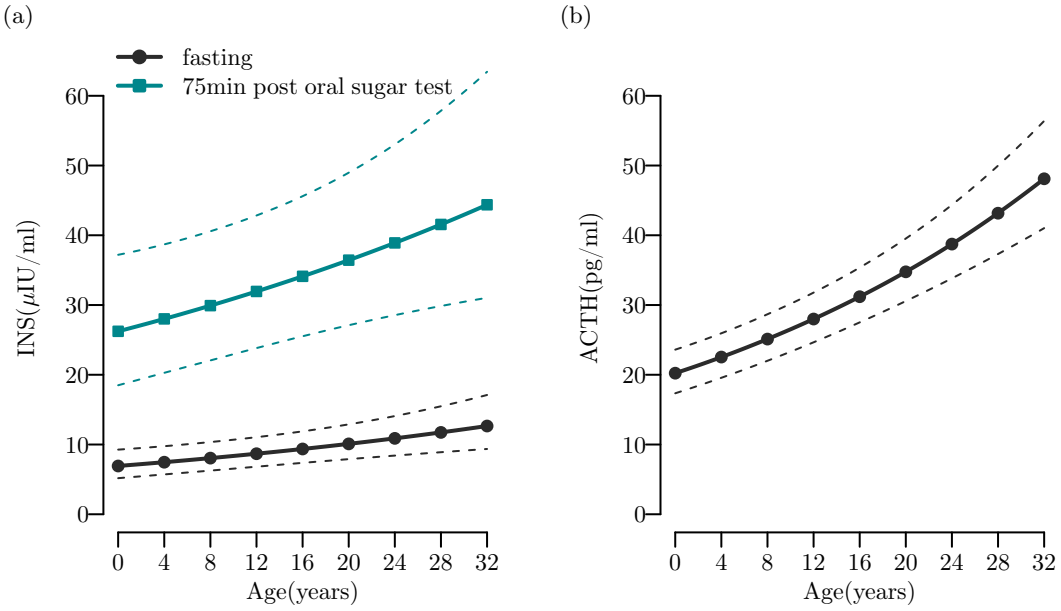
Different letters within a column indicate a statistically significant pairwise difference ($p<0.05$ with holm adjustment for multiple comparison).

Table 2.10: Predicted metabolic trait central means [95% confidence interval] by clinical group determined from the full multivariate, multilevel model

NH		**GH**		**GLU (mg/dl)**		**INS (μIU/ml)**	
ob$^-$lam$^-$	0.64[0.63,0.66]a	ob$^-$lam$^+$	1.18[1.17,1.20]a	ob$^-$lam$^-$	77.8[74.8,80.7]a	ob$^-$lam$^-$	5.5[4.3,7.0]a
ob$^-$lam$^+$	0.66[0.65,0.68]b	ob$^-$lam$^-$	1.19[1.17,1.20]a	ob$^+$lam$^-$	78.2[74.8,81.6]a	ob$^+$lam$^-$	8.4[6.3,11.2]b
ob$^+$lam$^-$	0.68[0.66,0.69]c	ob$^+$lam$^+$	1.23[1.21,1.25]b	ob$^+$lam$^+$	78.7[75.0,82.3]a	ob$^-$lam$^+$	9.6[7.3,12.5]b
ob$^+$lam$^+$	0.69[0.68,0.71]d	ob$^+$lam$^-$	1.23[1.21,1.25]b	ob$^-$lam$^+$	79.3[76.1,82.6]a	ob$^+$lam$^+$	13.6[10.0,18.4]c
GLU OST (mg/dl)		**INS OST (μIU/ml)**		**TG (mg/dl)**		**NEFA (mmol/L)**	
ob$^-$lam$^-$	99.2[93.4,105.0]a	ob$^-$lam$^-$	20.3[15.0,27.5]a	ob$^-$lam$^-$	24.2[20.8,28.2]a	ob$^+$lam$^-$	0.21[0.16,0.27]a
ob$^-$lam$^+$	101.0[94.6,107.3]a	ob$^+$lam$^-$	33.7[24.0,47.2]b	ob$^+$lam$^-$	27.3[23.0,32.6]a	ob$^-$lam$^-$	0.21[0.16,0.26]a
ob$^+$lam$^-$	102.2[95.8,108.6]a	ob$^-$lam$^+$	36.0[25.8,50.1]b	ob$^-$lam$^+$	34.1[28.9,40.3]b	ob$^-$lam$^+$	0.22[0.17,0.27]a
ob$^+$lam$^+$	102.6[95.6,109.6]a	ob$^+$lam$^+$	44.1[30.6,63.6]b	ob$^+$lam$^+$	34.6[28.7,41.7]b	ob$^+$lam$^+$	0.22[0.17,0.29]a
ACTH (pg/ml)		**LEP (ng/ml)**		**APN (ng/ml)**			
ob$^-$lam$^-$	27.6[24.2,31.4]a	ob$^-$lam$^+$	3.99[3.18,4.90]a	ob$^+$lam$^+$	2318[1669,3074]a		
ob$^+$lam$^+$	27.8[23.7,32.6]a	ob$^-$lam$^-$	4.23[3.47,5.06]a	ob$^-$lam$^+$	2402[1801,3089]a		
ob$^+$lam$^-$	28.9[24.9,33.6]a	ob$^+$lam$^+$	5.97[4.85,7.20]b	ob$^+$lam$^-$	3655[2871,4532]b		
ob$^-$lam$^+$	29.6[25.7,34.2]a	ob$^+$lam$^-$	6.22[5.15,7.38]b	ob$^-$lam$^-$	4061[3352,4839]b		

Different letters within a column indicate a statistically significant pairwise difference ($p<0.05$ with holm adjustment for multiple comparison). ob=obese, lam=history of prior laminitis

Figure 2.2: Predicted effect of age on the median trajectory of insulin(a) and ACTH(b). Dashed lines represent the 95% confidence interval. Estimate predictions were determined from the full model (estimates are adjusted for all variables included in the model).



2.4.5 Effect of environmental factors (diet, exercise, latitude, and month of year) on metabolic trait variation

The significance values from a Wald chi-square test of the overall effect of the explanatory variables dietary composition (Mcal, CP, NDF, Starch, and WSC per kg bwt per day), average hours per day grazing, average hours per day exercising, average hours per day stalled, and month at time of sampling on metabolic trait variation from the full model are presented in Table 2.11. Fixed effect estimates for diet, exercise, latitude, and month of year from the full model are reported in Appendix tables A.5-A.15. Additional covariates/explanatory variables in the full model included age, breed, sex, obesity and prior laminitis status, thyroxine supplementation (yes/no), and researcher collected sample vs owner submitted.

Figures 2.3 and 2.4 illustrate the relationship of significant environmental factors on metabolic trait variation. Daily dietary starch intake per kg bwt was negatively correlated with GLU OST levels. Hours per day spent in a stall was also negatively correlated with GLU OST levels. Hours per day spent grazing on pasture was positively correlated with NEFA levels while hours per day spent in a stall was negatively correlated with NEFA levels. ACTH levels correlated positively with latitude.

Metabolic traits that varied significantly with sampling month of the year are featured in Figure 2.5. GH was significantly lower in horses sampled during the month of July. Fasting GLU levels tended to rise throughout the spring and summer and then decrease during the fall and winter with GLU levels being significantly lower during the months of October, December, and January compared to the month of August. TG levels were significantly lower in June compared to February, May, August, September, and December The predicted effect of month on girth:height ratio, fasting glucose, triglyceride, ACTH, leptin, and adiponectin levels are illustrated in Figure 2.5. ACTH levels were also lowest during the month of June with significantly higher levels being found during the months of July through November. In addition, LEP levels were also lowest during the month of June with significantly higher levels observed during the months of February, May, and October. APN levels were lowest during the month of May.

Table 2.11: Significance results for association of environmental factors (Diet, exercise, latitude, and month p-values) with metabolic trait variation determined from the full multivariate, multilevel model

	+1 SD Mcal/kg bwt	+1 SD CP/kg bwt	+1 SD NDF/kg bwt	+1 SD Starch/kg bwt	+1 SD WSC/kg bwt	Hours grazing per day	Hours exercise per day	Hours stalled per day	Latitude	Month
NH	5.7e-01	5.5e-01	5.5e-01	4.2e-01	4.1e-01	5.3e-02	1.1e-01	2.8e-01	8.1e-01	3.2e-01
GH	9.5e-02	1.8e-01	9.5e-02	2.1e-01	2.4e-01	1.1e-01	7.5e-01	2.7e-01	1.5e-01	**2.6e-04**
GLU	8.1e-01	9.8e-01	6.6e-01	2.0e-01	1.5e-01	2.9e-01	3.5e-01	6.6e-01	6.5e-01	**9.9e-04**
logINS	9.8e-01	7.5e-01	2.3e-01	7.5e-01	8.8e-02	3.7e-01	6.7e-01	4.9e-01	3.4e-01	5.7e-01
GLU OST	4.9e-01	3.0e-01	2.4e-01	**3.7e-02**	3.3e-01	9.1e-01	7.3e-01	**1.1e-02**	8.5e-01	6.1e-01
logINS OST	5.0e-01	5.8e-01	7.0e-01	4.6e-01	2.4e-01	9.2e-01	8.2e-01	6.5e-01	7.1e-01	4.6e-01
logTG	1.7e-01	3.9e-01	4.5e-01	6.0e-01	5.0e-01	2.7e-01	8.8e-01	6.7e-01	4.3e-01	**3.6e-03**
$\surd$NEFA	9.8e-01	3.8e-01	6.1e-01	3.8e-01	9.1e-01	**1.3e-05**	9.4e-01	**6.7e-03**	1.6e-01	1.7e-01
logACTH	8.8e-01	5.0e-01	8.1e-01	7.4e-01	8.9e-01	9.4e-01	4.2e-01	1.7e-01	**2.6e-02**	**2.5e-14**
$\surd$Leptin	7.4e-01	5.3e-01	5.4e-01	3.1e-01	6.5e-01	9.0e-01	5.1e-01	5.7e-01	7.9e-01	**2.6e-10**
$\surd$Adiponectin	9.7e-01	7.6e-01	9.1e-01	6.5e-02	6.9e-01	2.1e-01	7.5e-02	7.0e-01	7.5e-01	**1.2e-02**

Figure 2.3: Predicted effects of starch consumption(a) and hours per day spent in stall(b) on median post-OST glucose and effects of hours per day grazing(c) and hours per day stalled(d) on median NEFA levels. Dashed lines represent the 95% confidence interval. Estimate predictions were determined from the full model (estimates are adjusted for all variables included in the model).

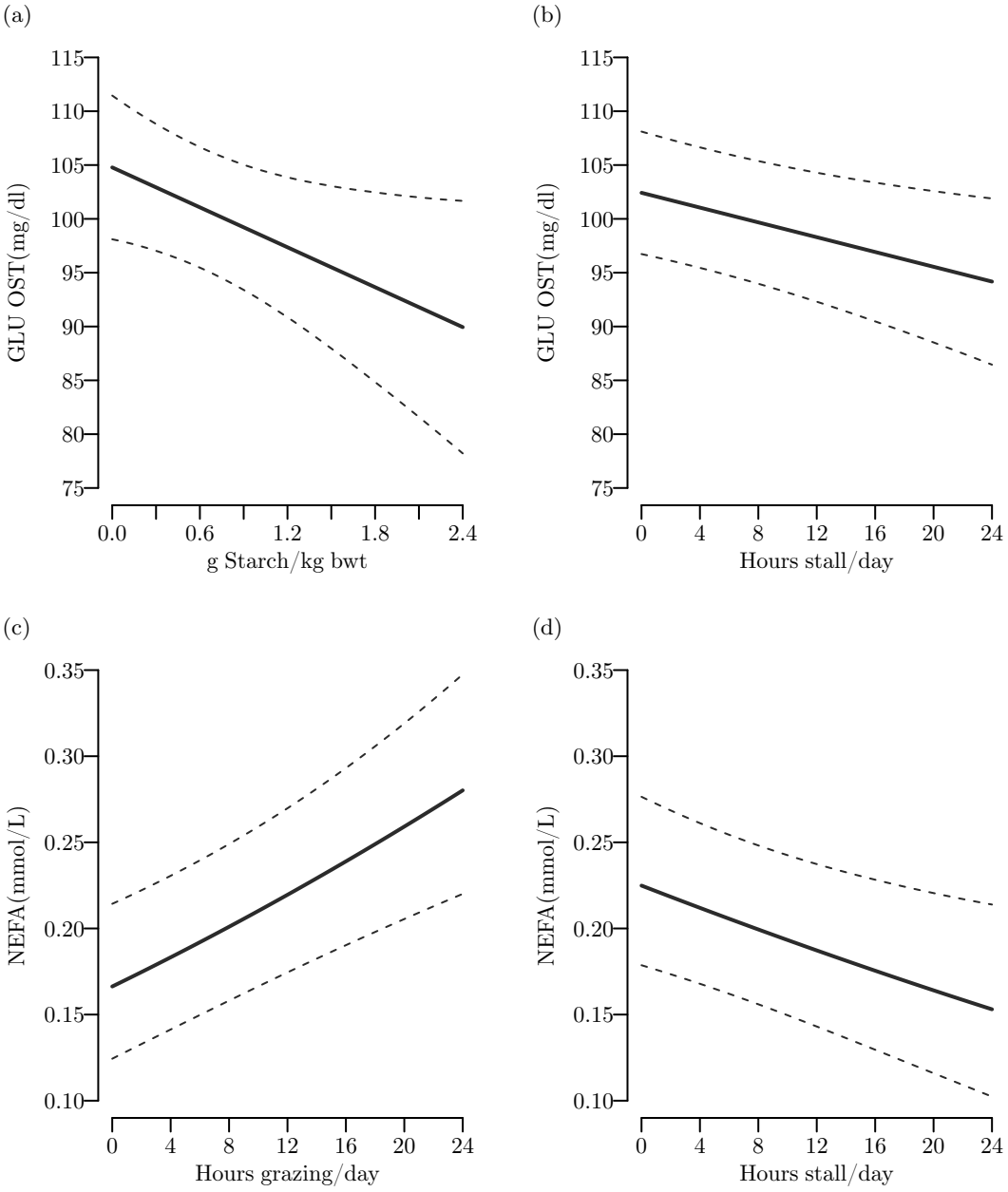


Figure 2.4: Predicted effect of latitude on ACTH levels. Dashed lines represent the 95% confidence interval. Estimate predictions were determined from the full model (estimates are adjusted for all variables included in the model).

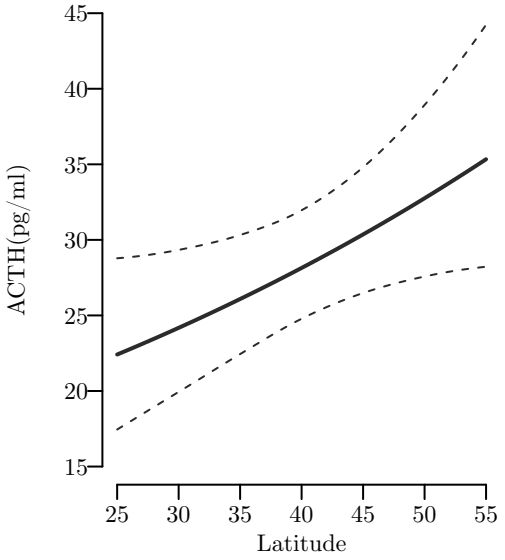
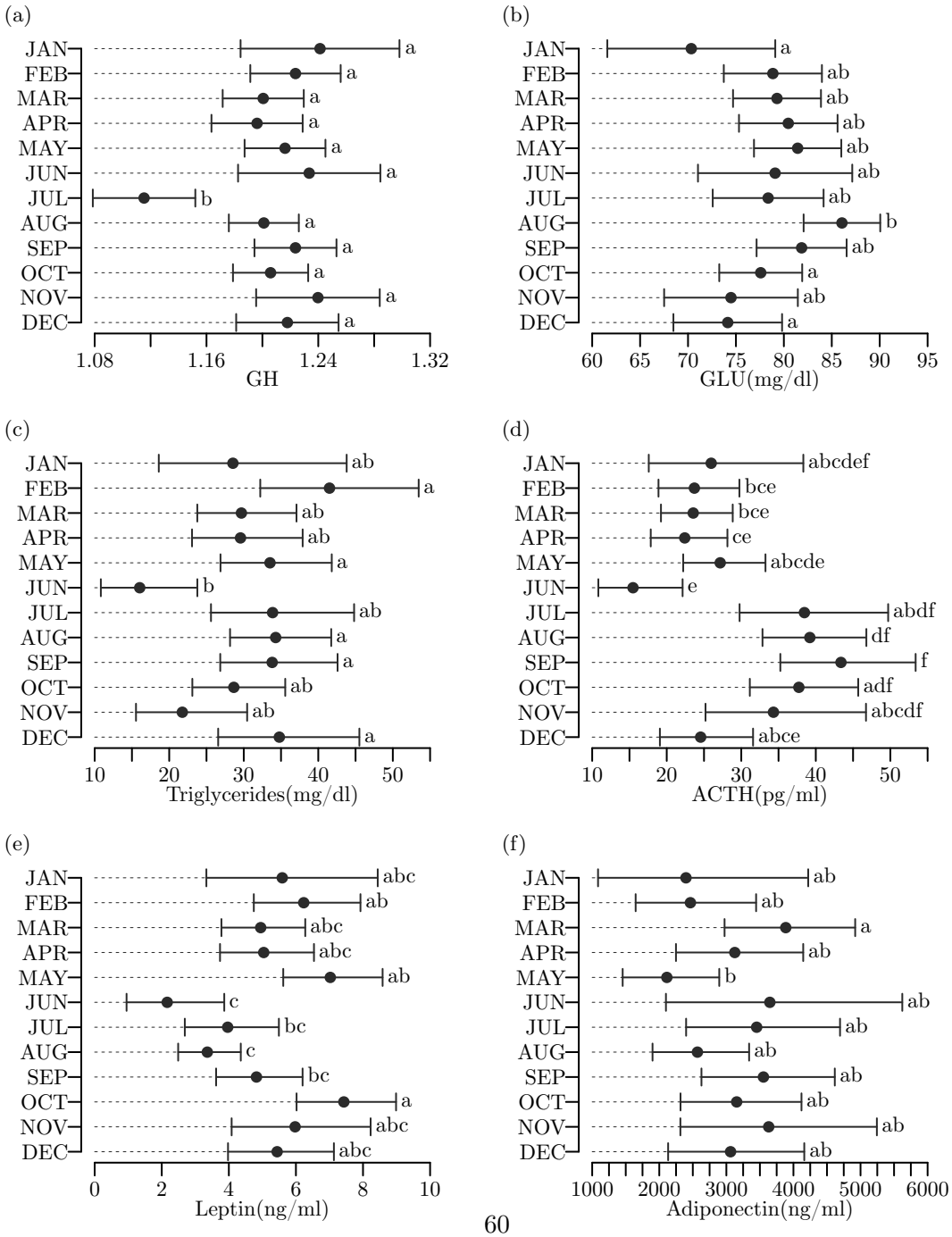


Figure 2.5: Predicted effect of month on girth:height ratio(a), fasting glucose(b), triglyceride(c), ACTH(d), leptin(e), and adiponectin(f) levels. Error bars represent the 95% confidence interval. Different letters indicate a statistically significant pairwise difference ($p<0.05$ with holm adjustment for multiple comparison). Estimate predictions determined from the full model (adjusted for all variables included in the model).



2.4.6 Relative importance of the full model predictor variables in explaining metabolic trait variation at the farm and individual horse level

The results presented in this section report both the percentage of metabolic variation at the farm and individual level that was explained by inclusion of the full model explanatory variables, in addition to the percent of variation that remained unexplained in the at the farm and individual level of the study population. These results provide an indication of the goodness of fit of the model, along with relative importance of the explanatory variables.

Figure 2.6 illustrates the variance decomposition of the morphometric traits NH and GH. In the null model, 47% of NH variance was present at the farm level. Inclusion of explanatory variables in the full model were unable to explain any of the variance at the farm level. Approximately 25% of the individual level variance was able to be explained by the model with obesity status accounting for approximately half of the explained NH variance at the individual level. In the null model, 43% of GH variance was present at the farm level. Approximately 28% of the farm level variance was able to be explained with sampling month accounting for half of the explained variance. At the individual level, similar to NH variance, obesity status contributed the largest amount to the explained variance. Sex was the second largest contributor to explained variance at the individual level for both NH and GH variance.

Variance decomposition of GLU (fasted and OST) is presented in Figure 2.7. In the null model, 35% of fasted GLU variance was present at the farm level. Sampling month of year was the largest contributor to explained variation at the farm level. Approximately 80% of fasted GLU variance at the farm level remained unexplained and variance at the individual level remained largely unexplained (96%). 42% of variance in post-OST GLU was present at the farm level. Dietary components, especially starch intake, were important factors in explaining post-OST GLU variance at the farm level. Although 87% of variance at the farm level remained unexplained and an even higher percentage remained unexplained at the individual level (97%).

Variance decomposition of INS (fasted and OST) is presented in Figure 2.8. In the null model, 36% of fasted INS variance was present at the farm level. Approximately 25% of variance at the farm level was explained by factors in the full model. Obesity

status, breed, and prior laminitis status were important factors in explaining fasted INS variance at both the farm and individual level. 77% of fasted INS variance at the farm level remained unexplained. 33% of post-OST variance was present at the farm level. Breed was the most important factor in explaining post-OST variance at the farm level, while prior laminitis and obesity status were the most important factors explaining post-variance at the individual level. About 87% of farm level and 76% of individual level post-OST variance remained unexplained.

TG and NEFA variance decomposition is presented in Figure 2.9. 24% of TG variance was present at the farm level. Prior laminitis status was the most important factor in explaining TG variance at both farm and individual level. 74% of farm level and 83% of TG variance at the individual level remained unexplained. 44% of NEFA variance was present at the farm level. Hours per day spent grazing accounted for 11% of farm level NEFA variance. The vast majority of NEFA variance at the individual level remained unexplained (97%).

A large percentage of ACTH variance occurred at the farm level (49%) (Figure 2.10). Sampling time of year explained 40% of ACTH variance at the farm level. Age explained 10% of variance at the individual level.

The variance decomposition of adipokines LEP and APN is illustrated in Figure 2.11. 31% of LEP variance was present at the farm level with sampling month and breed explaining 31% and 22% of LEP variance at the farm level, respectively. Obesity status was the most important factor in explaining LEP variance at the individual level. 23% of APN variance was present at the farm level. Samples collected by a farm's veterinarian and submitted by the farm owner instead of a researcher visiting the farm and collecting the sample explained 20% of APN variance at the farm level. Breed and sampling time of year were also important in explaining farm level APN variance. Prior laminitis status was the most important factor in explaining APN variance at the individual level.

Total explained variance and explanatory variable contribution to the total variation is reported in Table 2.12. From Figure 2.11a it can be determined that obesity status accounts for 7.4% of the total variation in leptin levels by multiplying the Null farm and horse level variance estimates by the percent contribution of the factor to the farm and horse level variance, respectively, and summing the products (e.g.

$(31.5 * 0.039) + (68.5 * 0.09) = 7.4)$.

Figure 2.6: Variance decomposition of morphometric traits neck circumference:height ratio(a) and girth:height ratio(b)

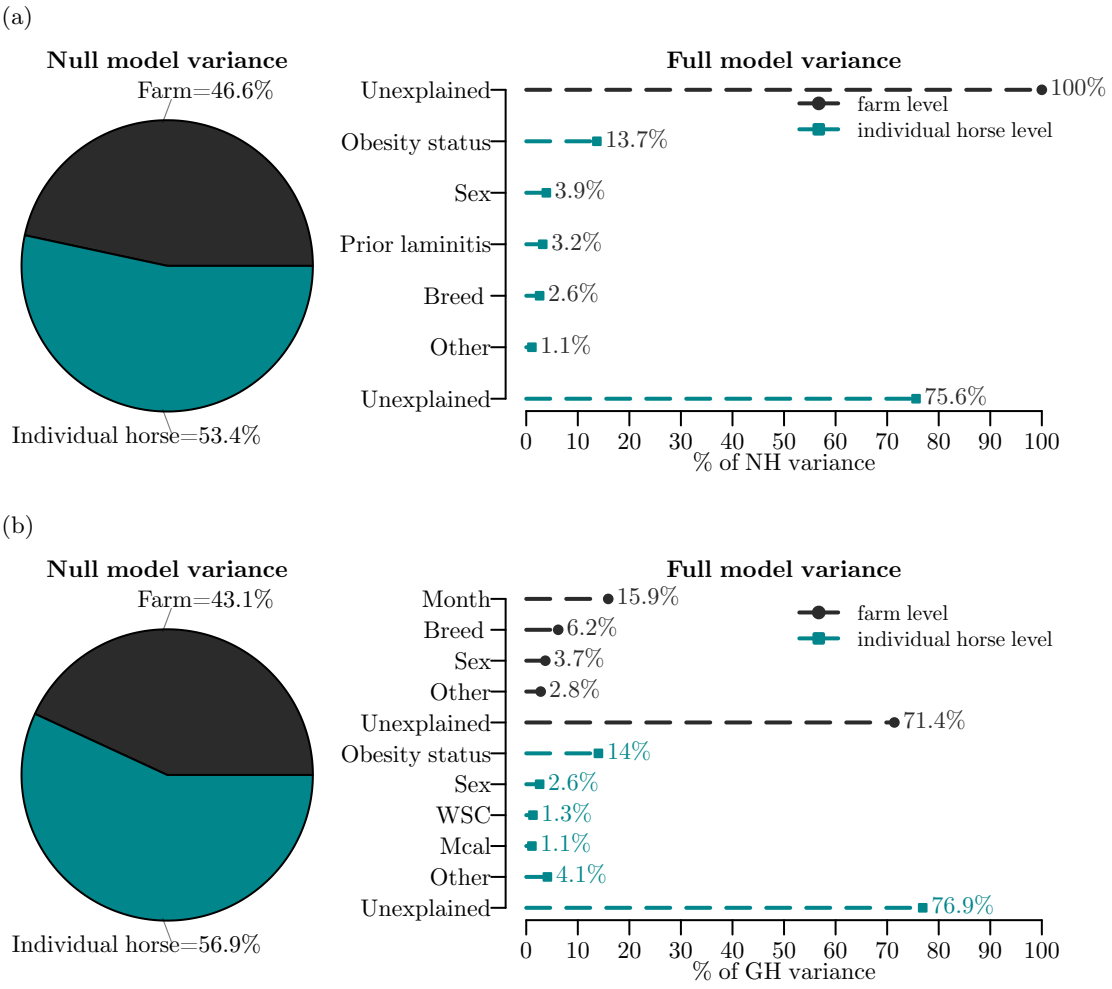
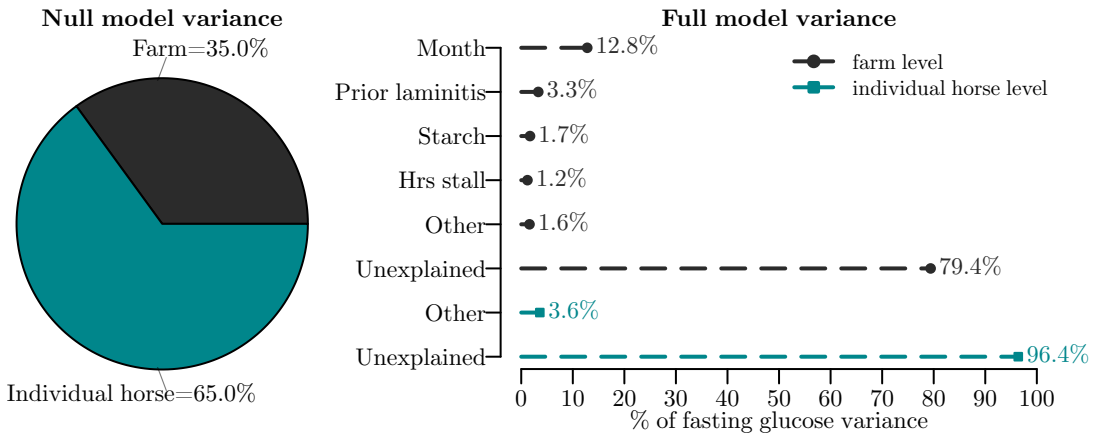


Figure 2.7: Fasting(a) and post-OST(b) glucose variance decomposition

(a)



(b)

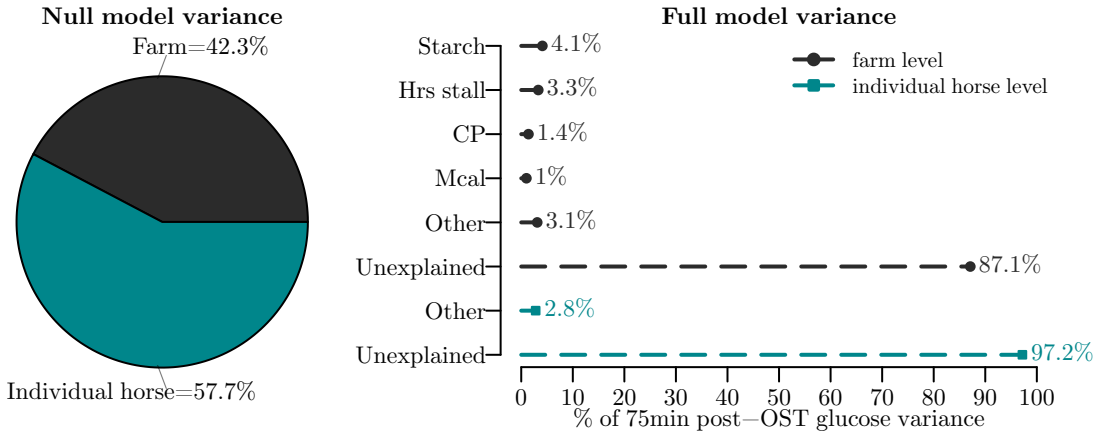
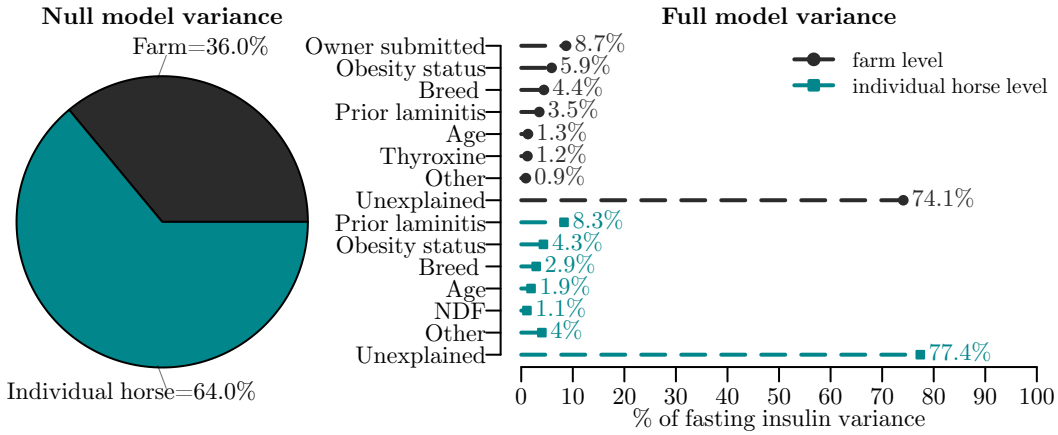


Figure 2.8: Fasting(a) and post-OST(b) insulin variance decomposition

(a)



(b)

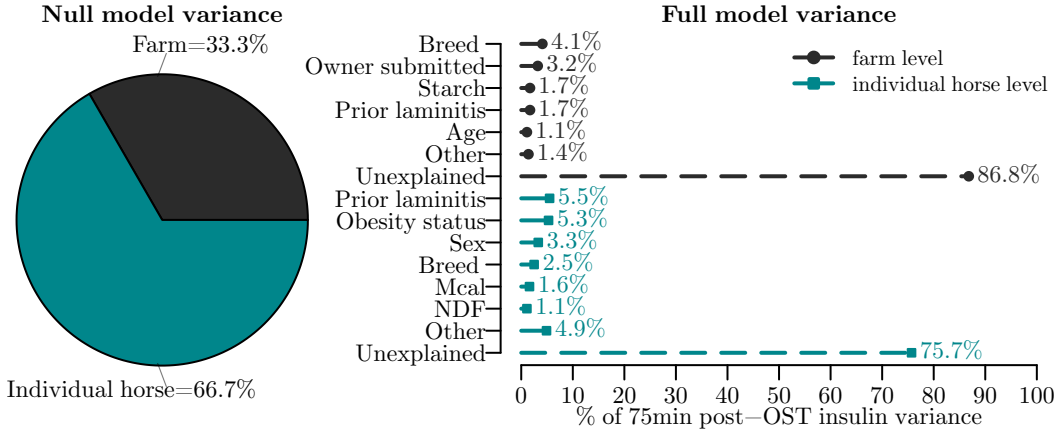


Figure 2.9: Variance decomposition of triglycerides(a) and NEFA(b)

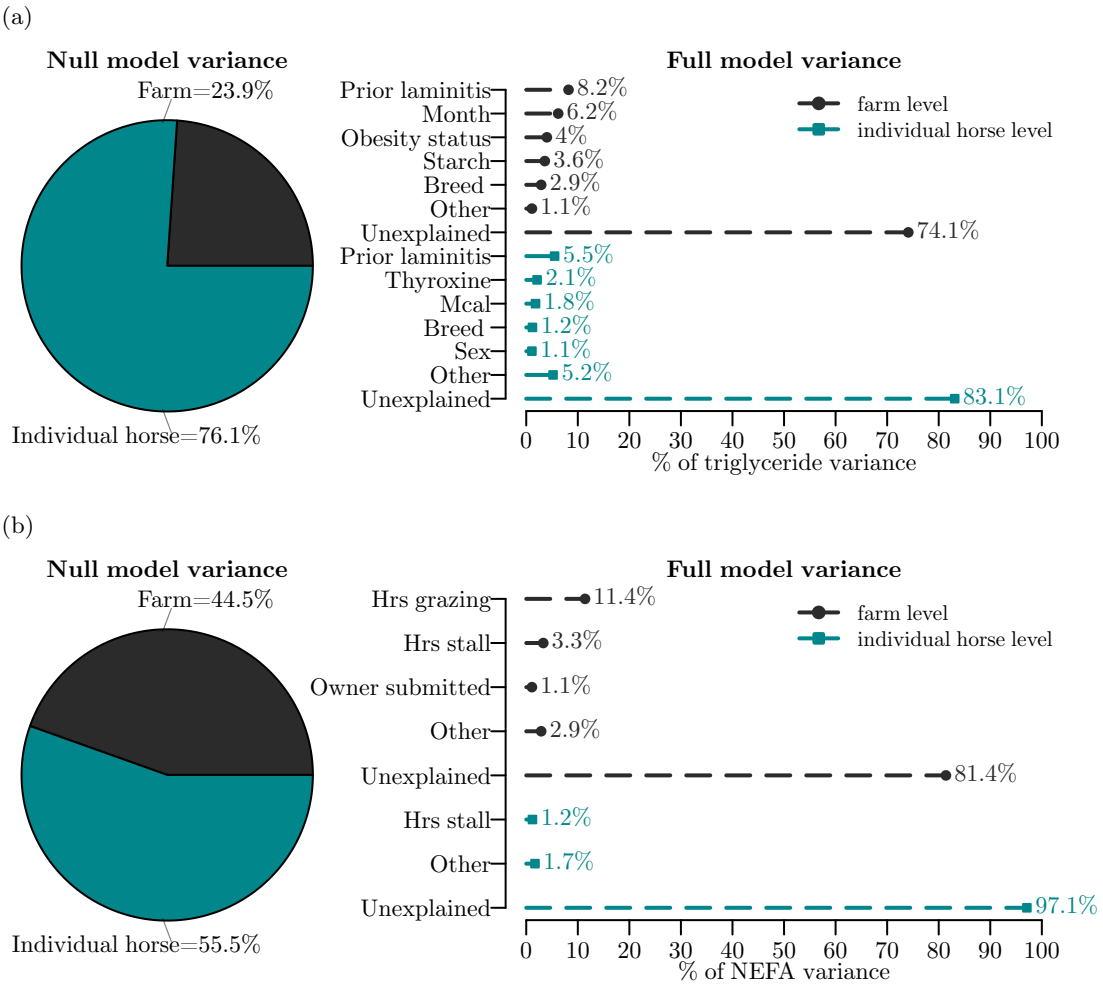


Figure 2.10: Variance decomposition of ACTH

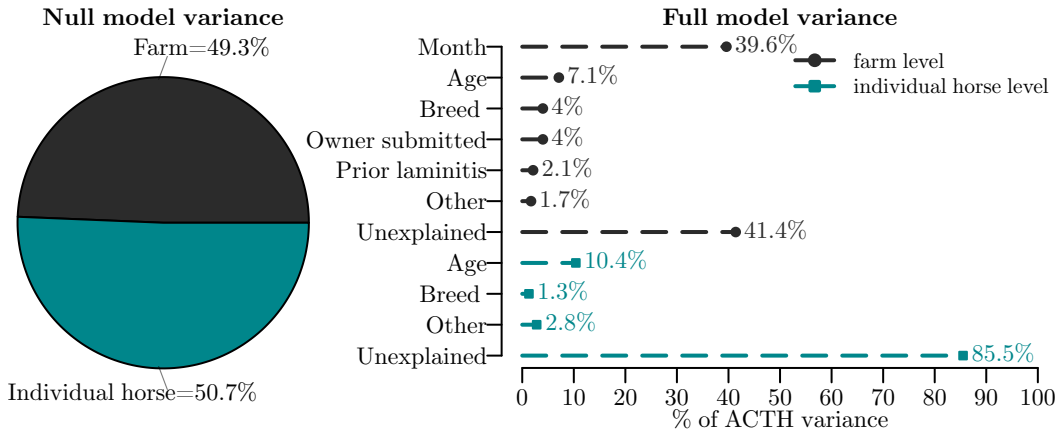
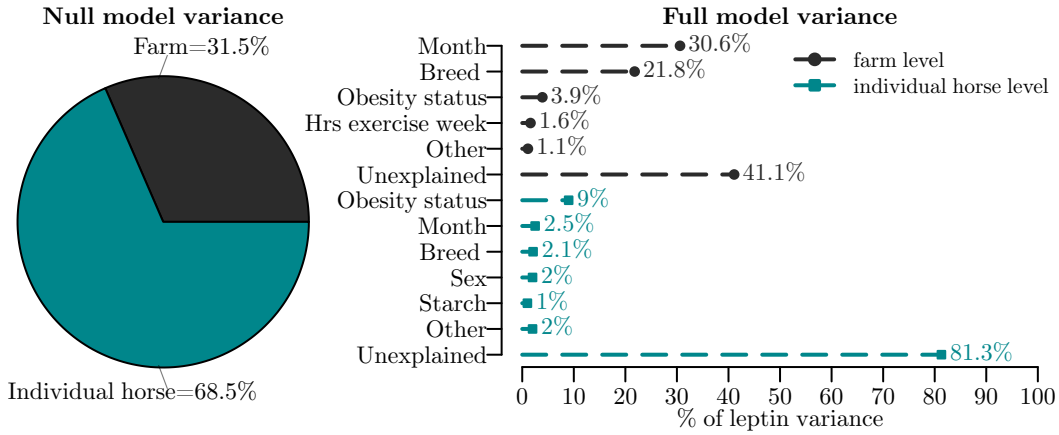


Figure 2.11: Variance decomposition of adipokines leptin(a) and adiponectin(b)

(a)



(b)

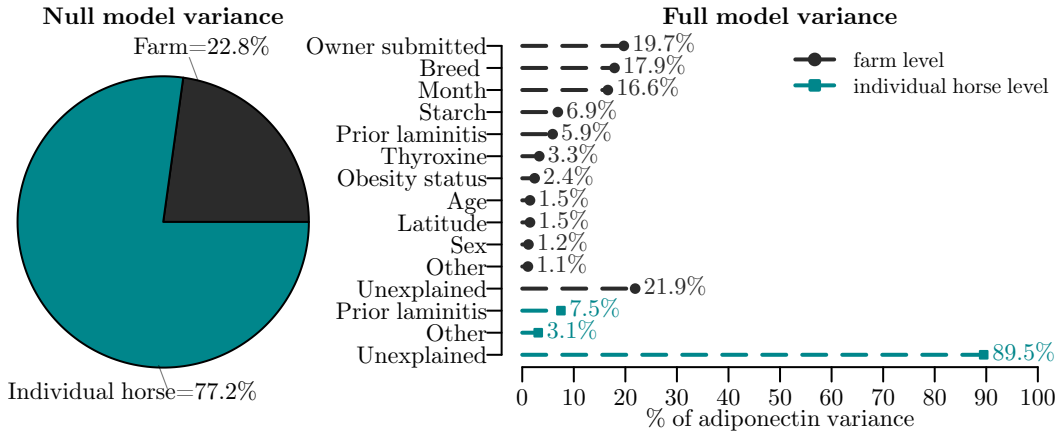


Table 2.12: Explanatory variable contribution to total variation in metabolic traits

Parameter	NH	GH	GLU	INS	GLU OST	INS OST	TG	NEFA	ACTH	LEP	APN
Breed	1.4%	3.1%	0.5%	3.4%	0.3%	3.0%	1.6%	0.3%	2.6%	8.3%	4.1%
Sex	2.1%	3.1%	0.7%	0.4%	0.5%	2.2%	0.9%	0.4%	0.3%	1.3%	0.4%
Age		0.2%		1.7%	0.3%	1.0%	0.4%		8.8%		0.3%
Laminitis	1.7%		1.1%	6.6%	0.4%	4.2%	6.2%	0.1%	1.1%	0.2%	7.1%
Obesity	7.3%	8.0%	0.1%	4.9%	0.2%	3.8%	1.4%	0.2%	0.3%	7.4%	0.8%
Month		6.9%	4.6%	0.5%	0.1%	0.6%	2.1%	0.4%	19.6%	11.4%	3.8%
Latitude		0.3%		0.1%	0.1%		0.1%		0.3%		0.3%
Hrs in stall	0.1%	0.3%	0.4%	0.3%	1.4%	0.1%	0.1%	2.1%	0.4%		
Hrs exercise		0.3%		0.2%		0.1%	0.1%	0.1%	0.3%	0.5%	0.2%
Hrs grazing		0.6%	0.2%		0.1%	0.2%		5.4%			
Mcal	0.1%	0.6%	0.2%	0.4%	0.5%	1.1%	1.3%	0.1%	0.2%	0.6%	0.6%
CP	0.1%	0.4%	0.1%	0.3%	0.6%	0.7%	0.7%		0.2%	0.4%	0.3%
NDF		0.4%	0.2%	0.7%	0.3%	0.7%	0.6%	0.1%	0.1%	0.2%	0.2%
Starch	0.1%	0.3%	0.9%	0.1%	2.0%	0.8%	1.5%	0.1%	0.1%	0.7%	2.0%
WSC	0.1%	0.8%	0.2%	0.3%	0.1%	0.6%	0.4%		0.1%	0.2%	0.1%
Owner submitted		0.1%	0.1%	3.2%		1.1%		0.5%	2.0%	0.1%	4.5%
Thyro-L supplementation	0.1%	0.4%	0.1%	0.7%	0.2%	0.3%	1.6%				1.1%
Total explained variance	13.0%	25.5%	9.6%	23.8%	7.1%	20.6%	19.1%	9.9%	36.3%	31.4%	26.0%

2.5 Discussion

Although equine metabolic syndrome (EMS) is a term frequently used by veterinary clinicians, there is a lack of consensus regarding the diagnostic criteria that define EMS. Generalized/regional adiposity, hyperinsulinemia, insulin resistance, dyslipidemia, and predisposition to laminitis development are reported components of the syndrome. However, not all studies have consistently reported the same phenotypic components for EMS, which may suggest different confounders contributed to the phenotypic variation among the differing sample cohorts and experimental designs. For example, Measurements to quantify adiposity, GH and NH, have also been associated with the EMS phenotype in horses[79] and proposed as predictors of laminitis risk in ponies;[78] however obesity has not been a consistent finding, particularly in well-managed populations.[81] Hyperinsulinemia has been shown to be a feature of the EMS phenotype although the values vary between studies.[78,81,115] Elevated TG have been associated with the EMS phenotype in ponies[78,81,116] but not in horses;[79] whereas, NEFA concentrations have been associated with an EMS phenotype in horses,[79] but were not useful in the differentiation of an EMS phenotype in ponies.[77,78] The shortfalls of prior EMS studies include small study populations, studies conducted in a single breed, and/or studies conducted at a single geographic location, which may prevent translation of findings to the larger population. Therefore the overarching objective of this study was to investigate variation and co-variation in metabolic traits in a large population of horses and ponies over a range of individual and environmental conditions.

2.5.1 Equine metabolic traits cluster at the farm level

A multilevel, multivariate approach made it possible to 1) determine the degree to which a metabolic trait clusters in horses sampled from the same farm, 2) assess the correlation of different metabolic traits at both a farm and individual level and 3) determine the effects of individual and farm level characteristics on metabolic trait variation. All of the metabolic traits included in the study exhibited substantial clustering at the farm level with percentage of variance present at the farm level ranging from 23-49% of the total trait variance (Table 2.4). Clustering of traits at the farm level indicate factors

shared by horses sampled from the same farm contribute to variation in metabolic traits.

2.5.2 Equine metabolic traits correlate at both the farm and individual horse level

As expected GLU and INS levels (fasted and post-OST) exhibited a strong, positive correlation. Interestingly, the correlation was very strong at the farm level indicating farm related factors play a role in the correlation between GLU and INS measurements. Inclusion of explanatory variables in the full model revealed farm characteristics such as sampling time of year, hours spent in a stall and diet (starch in particuular) as important factors effecting GLU variation which could in turn explain the strong GLU and INS correlation at the farm level.

At the individual level, correlation of metabolic traits were similar to that observed in humans[132] with positive correlations among INS, LEP, and TG levels and a negative correlation of APN with the aforementioned traits. A measure of regional adiposity (NH), more so than a measure of generalized adiposity (GH), correlated with INS, TG and APN levels while LEP exhibited a stronger correlation with regional adiposity (GH). Again, these findings mirror relationships of different fat depots with varying biochemical profiles observed in humans.[132] Previous data suggest the nuchal ligament adipose tissue is more likely than other adipose depots to display an inflammatory phenotype in the horse and therefore may play a unique role in the pathogenesis of metabolic dysregulation,[133] observing correlation of regional adiposity with biochemical measures associated with prior laminitis status (INS, TG, APN) supports this hypothesis. In humans, subcutaneous fat in the neck region has been found to be strongly associated with insulin resistance.[134]

Some traits tend to correlate more strongly at the farm level but not at the individual level. For example, NEFA and post-oral sugar test glucose (GLU OST) were correlated at the farm level, indicating farm level characteristics may explain variation in energy production from fat oxidation versus glucose utilization. Also, ACTH and APN were also correlated at the farm level but not the individual level. From the Shapley value analysis, it was able to be determined that season was a large contributor to the farm level variance in both of these traits.

2.5.3 Clinical group effects metabolic trait measures

Clinical groups based on obesity and prior laminitis status demonstrated divergent patterning of metabolic traits. Elevated TG levels were significantly higher in individuals with prior laminitis, but were not significantly increased in individuals that were obese without a history of laminitis. Whereas both obesity and prior laminitis impact INS and INS OST levels. GH, a measurement of regional adiposity, was only associated with obesity status and did not vary with prior laminitis status, as opposed to NH, a measure of regional adiposity, that was greater in both obese horses and those with prior laminitis. Adipokines LEP and APN also exhibited divergent patterns according to clinical group status. LEP levels were elevated in obese individuals to a similar degree in horses with and without a prior history of laminitis. APN levels were decreased in horses with a prior history of laminitis to a similar degree in non-obese and obese individuals. These findings suggest LEP may be a good indicator of total body fat mass in horses and decreased APN may indicate the presence of an unhealthy state that predisposes horse to laminitis. APN and LEP have been identified as key components of human metabolic syndrome. Interestingly, APN has been identified as a better predictor than LEP of the presence and degree of coronary atherosclerosis, independent of body mass index.[135]

2.5.4 Physiologic factors influence equine metabolic trait variation

Breed differences were observed for several traits (NH, GH, INS (fasted and OST), ACTH, LEP, APN). The QH breed was the most divergent with lower INS/INS OST, LEP, and higher APN. Increased muscle mass in the QH breed provides a possible explanation for observing a more insulin sensitive phenotype. A variant involving the myostatin gene occurs frequently in the QH breed and is associated muscle fiber type proportions.[136] In the mouse, myostatin deficiency has been linked with reduced diet-induced obesity and a more insulin sensitive phenotype.[137–142] Other breed differences included observing higher ACTH levels in the pony breeds, a finding that has not been identified in previous studies.[143,144] LEP levels were highest in the Morgan breed. LEP levels are known to be correlated with total body fat mass in humans[145] and therefore may indicate higher total body fat mass in the Morgan breed. Breed differences suggest

genetic differences as a source of metabolic trait variation and warrants further investigation of genetic sources of trait variation. Although a drawback of observing breed differences for metabolic traits is the limitations it poses on the use of a single reference range for all breeds when utilizing metabolic traits as a diagnostic.

Gender differences were observed for several traits with mares having a significantly lower measure of regional adiposity (NH) and stallions having a lower measure of generalized adiposity (GH), suggesting gender differences in equine body fat distribution. This study included a substantial number of stallions (n=32) which aided in the ability to detect significantly higher insulinemic responses to an oral glucose challenge than both mares and geldings. Mares also had higher LEP levels than geldings. Gender differences in triglyceride levels have been reported to correlate with fat deposition sex differences in humans.[146] Gender differences in LEP levels in humans are accounted for by percent body fat.[147] Similar to previous reports,[144,148] age was positively associated with both fasting and OST INS and to a larger degree with ACTH levels.

2.5.5 Environmental factors influence equine metabolic trait variation

Seasonal variation in metabolic traits observed in our study cohort were similar to what has been previously reported, despite the fact that samples were obtained from each horse on only a single day of the year. Elevated ACTH in the fall months has been previously reported in horses.[149–151] However, we report a positive correlation between latitude on ACTH levels, which is opposite to the relationship observed previously in horses[143] but consistent with the concept of higher latitudes experiencing greater seasonal extremes and thus having larger photoperiod effects. In addition to photoperiod, differences in ambient temperature may also potentially explain higher ACTH values in horses from higher latitudes.[152]

Seasonal variation in GLU levels have been documented in other mammalian species[153] with glucose levels reported to be highest in the fall in humans[154] similar to our findings. Frank et al.[150] also reported higher GLU levels for horses during the month of September, in addition to elevated insulin concentrations. Similar to Place et al.,[149] the effect of month on insulin levels was not significant in this study, although Schreiber et al.[151] reported insulin levels to be higher in older horses during the month of September.

Borer et al.[155] observed a significant season prior laminitis interaction for insulin levels. Age x month and age x clinical group interactions were not examined in the current study but may warrant further investigation.

GH and LEP were lower during the summer months. A possible explanation for this observation may be seasonal changes in body fat mass. Horses are long day breeders and exhibit photoperiod responsive changes in hormone levels. Ferrets and mink are also long day breeders and exhibit bodyweight gain during short days and weight loss during long days.[156,157] Therefore it is plausible weight loss in the summer months independent of diet and exercise may be due to changes in photoperiod and similar biologic mechanisms may be responsible in both species. Similar to previous reports, leptin values were highest in the month of October, at the end of the ACTH rise.[158] It is possible increased ACTH levels may have induced leptin resistance in an effort to prepare for winter.[159] The study findings suggest seasonality is an important source of metabolic trait variation; further evidence of this conclusion is the substantial reduction in farm level variance achieved for several traits with the inclusion of month as an explanatory variable. The horse, similar to other long day breeders, seems to switch to an insulin resistant state during the fall. This "switch" would allow a horse to increase fat stores to be utilized throughout the winter when food sources may be limited and increased energy is needed for thermoregulation and to support development of a fetus in pregnant mares. Increased insulinemic responses have been reported in pregnant mares compared to non-pregnant mares during an intravenous glucose tolerance test.[160]

Minimal effects for diet and exercise parameters on metabolic trait variation were detected. Dietary starch intake demonstrated a negative correlation with post-OST GLU levels. A reduced glycemic response to an oral sugar test has previously been reported both in horses adapted to a high starch diet,[161] similarly, humans are instructed to consume a higher carbohydrate diet a few day prior to an oral glucose tolerance test in order to prevent false positives (especially individuals accustomed to a low carb diet).[162] Hours per day grazing correlated positively with NEFA levels where as hours per day spent in a stall correlated negatively with NEFA levels. NEFA levels are influenced by the availability of food and tend to peak before the next meal. Lower NEFA levels would be expected in horses that spent more time grazing, however it is possible that these horses were less acclimated to the fasting period they experienced prior to blood

sample collection resulting in higher NEFA levels. Testing NEFA levels from non-fasted samples would be necessary to confirm this hypothesis. Correlations of metabolic traits with hours of exercise per week were not detected. Power may have been limited to detect an association with exercise due to only 15% of the population having received more than 3 hours of exercise per week.

2.5.6 Unexplained equine metabolic trait variation remains at both the farm and individual horse level

Although a moderate amount of variation in metabolic traits was explained by the explanatory variables included in the full model, the substantial amount or remaining unexplained variation suggest additional unaccounted for sources of farm and individual level variance may exist. It is likely that individual genetic differences exist which contributes to metabolic trait variation. The detection of breed differences in our study support this hypothesis. Identification of genetic variants associated with metabolic traits and subsequent genotyping of the individuals in this study cohort for those genetic variants might potentially explain additional metabolic trait variance. Additional environmental factors not measured in this study may also potentially impact metabolic trait variation. For example, chemicals present in the environment have endocrine disrupting capabilities and could be a potential unaccounted for source of variation in equine metabolic traits.[163]

2.5.7 Diagnostic implications of individual and farm level sources of metabolic trait variation

The present study identified numerous factors beyond obesity and prior laminitis status contribute to variation in metabolic traits, including seasonal effects, breed, age, sex, and diet. These findings obtained by examining equine metabolic trait variation in a large population of horses under varying environmental and individual conditions provide an explanation for the discrepancy in some conclusions drawn from previous studies regarding the EMS phenotype. However, our findings also draw attention to

the difficulty in determining appropriate reference ranges for EMS diagnostic criteria in the presence of multiple sources of "normal" metabolic variation. A potential adjunct for improving EMS diagnosis prior to laminitis development and monitoring a horse's response to management would be to test multiple horses on the same property. A horse with values dissimilar from the rest of the herd may be at higher risk of laminitis development. Monitoring additional horses may be even more useful for monitoring progression of EMS. Instead of only monitoring the affected horse and questioning if the horse's values changed due to individual factors (progression/improvement of EMS) or due to environmental factors, one could monitor additional horses to determine if all of the horses experienced a similar change in test results due to farm level factors (season, diet).

2.5.8 Study limitations and future directions

A limitation of the present study is the one-time sampling of the study population. The present study was not designed to ask questions such as: does metabolic trait variation and co-variation in previously laminitic horses vary with changes in bodyweight? This should be a focus area of future research along with identifying additional sources of variation in metabolic traits in horses with a prior history of laminitis under obese and non-obese conditions. Further research is needed to identify additional environmental sources of metabolic trait variation and to dissect sources of individual level metabolic trait variation. Breed differences were identified by the present study however further work is necessary to determine genetic variants responsible for breed variation in metabolic traits.

In summary, our results indicate multiple sources of variation in equine metabolic trait variation, in addition to differences in metabolic trait variance and covariance dependent on obesity and prior laminitis status. Given the strong association of adiponectin with prior laminitis status, incorporation of hypoadiponectinemia into the definition of EMS should be considered.

2.5.9 Main findings of Chapter 2 ("Take home messages")

1. Equine metabolic trait variation is associated with both measured (i.e. age, sex, breed, obesity status, and prior laminitis status) and unmeasured individual level factors in addition to measured (i.e. season) and unmeasured environmental level factors.

2. Obesity and prior laminitis status are associated with divergent metabolic profiles. Elevated blood leptin concentration is associated with obesity status where as decreased adiponectin and increased triglyceride concentration are associated with prior laminitis status. Thus, blood concentrations of leptin and insulin alone (insulin is associated with both obesity and prior laminitis status) may not be ideal biomarkers of laminitis risk. Adiponectin and triglyceride blood concentrations may be more appropriate biomarkers of laminitis risk. Moreover, obesity was not found to be essential in expressing a "metabolically unhealthy" phenotype characterized by low adiponectin and elevated triglyceride and insulin concentrations; suggesting lean horses may also be at risk for developing endocrinopathic laminitis.

3. Individual level factors, other than obesity and prior laminitis status (i.e. age, sex, breed), and environmental level factors (i.e. season) are also associated with metabolic trait variation and thus present challenges for determining appropriate diagnostic reference ranges.

4. Undetermined sources of metabolic trait variation at both the individual and environmental level in horses remains. Potential sources of individual (e.g. genetic variants) level variation and environmental (e.g. organic pollutants) level variation warrant investigation.

Chapter 3

Characterization of the equine response to an oral sugar challenge

3.1 Summary

Differences in the incretin response may also contribute to equine metabolic trait variation. The incretin response, defined as the difference in the insulinemic responses between an oral and intravenous glucose challenge, is controlled by intestinal secretion of peptides, such as GLP-1, that stimulate pancreatic insulin secretion. While the incretin response has been hypothesized to play a role in the EMS pathogenesis, this hypothesis has not been adequately tested. In Chapter 3, the glycemic, insulinemic, and total and active GLP-1 responses to an oral sugar challenge, and the activity of DPP4, the major protease that breaks down GLP-1, were characterized. The use of a longitudinal analysis, rather than the traditional area under the curve analysis, allowed for increased power to detect differences in these responses, including variation due to breed, obesity, and prior laminitis status.

3.2 Introduction

Hyperinsulinemia and insulin resistance are components of equine metabolic syndrome (EMS), a condition associated with predisposition for laminitis development.[76] Incretin hormones play a key role in glucose homeostasis and the insulin secretory response. The incretin hormones glucagon-like peptide-1 (GLP-1) and glucose-dependent insulinotropic polypepide (GIP) are produced by endocrine cells of the gastrointestinal tract and promote the release of insulin in response to nutrient ingestion.[164] Both GLP-1 and GIP have a short half-life due to rapid inactivation by the enzyme dipeptidyl peptidase-4 (DPP4).[165] Impairment of the incretin response has been implicated in human obesity and type 2 diabetes.[166] Decreased GLP-1 baseline measurements and responses to oral glucose have been associated with oral glucose intolerance.[167,168] Both GLP-1 analogs and DPP4 inhibitors are currently used therapeutics for the management of type 2 diabetes.[169] Alteration of the equine incretin response has been hypothesized to play a role in the pathogenesis of EMS, however limited research has been performed in terms of testing this hypothesis.[92] The objectives of the current study were to 1) dynamically model the equine response to an oral sugar challenge and 2) identify physiologic and genetic factors associated with variation in this response.

Candidate genes dipeptidyl peptidase-4 (*DPP4*) and proglucagon (*GCG*) were selected for investigation in this study based on their function, in addition to localizing to a genomic signature of selection region on chromosome 18 in the Morgan horse.[170] The proglucagon gene (*GCG*) encodes glucagon and several glucagon-like peptides (including GLP-1) that are liberated in a tissue specific manner in pancreas, intestine, and brain. The structure of the proglucagon gene is depicted in Figure 3.1. Pancreatic islet proglucagon transcription is regulated by enhancer-like elements in the promoter region.[171] Less is known about control of proglucagon transcriptional regulation in the endocrine gut cells. Based on available rat and human data, a proglucagon gene intestinal enhancer region, designated GUE, has been estimated to reside between -1252 and -1600 base pairs from the transcriptional start site.[171] Alignment of the equine reference sequence for the 5' flanking region with other mammalian species revealed an ~500 base pair insertion just upstream of the promoter region (Figure B.1). The region was identified by RepeatMasker[172] as an endogenous retrovirus class I long terminal repeat13b

(ERV class I LTR13b). Thus it is an important region to investigate given ERV LTRs naturally contain abundant transcriptional regulatory signals and often contribute to variation in mammalian gene expression, including tissue-specific expression, through donation of alternative promoters and enhancers.

A multilevel, multivariate longitudinal modeling approach was performed to characterize the variation and covariation of glucose, insulin, and GLP-1 trajectories in 65 Morgan horses and 65 Welsh ponies over a 120 minute time course based on measurements at 0,15, 30, 60, 75, 90, and 120 minutes following the oral sugar challenge. Factors including age, breed, gender, obesity/prior laminitis status, basal biochemical measurements (triglycerides, NEFA, leptin, adiponectin, DPP4-activity), dietary measurements (Mcal, CP, NDF, Starch, WSC), and polymorphisms within candidate gene regions were tested for association with variation of the mean curve trajectory (value at time zero, initial rate of change, and deceleration rate), in addition to single summary indices of the curve trajectory, area under the curve. It is common practice to use single summary measures of oral sugar tests with repeated measurements, however different OST curves can yield similar single summary measures and information of biological interest may be lost.[173,174] Therefore, investigation of the trajectories was incorporated in an effort to characterize the mean GLP-1, insulin, and glucose curves, their covariation, and identify factors associated with variation of the mean curve.

3.3 Methods

3.3.1 Sampling design

Glucose (GLU), insulin (INS), total GLP-1 (GLP1tot), and active GLP-1 (GLP1a) concentrations were determined in 56 Morgan horses and 56 Welsh Ponies from 7 farms at 0,15, 30, 60, 75, 90, and 120 minutes following an oral sugar challenge (0.15 ml corn syrup per kg bodyweight). Triglyceride (TG), non-esterified fatty acid (NEFA), leptin (LEP), adiponectin (APN), and DPP4 activity blood concentrations were determined at the zero-minutes time point prior to oral sugar administration. Horses were removed from pasture the evening prior to the AM blood sample collection. Horses were allowed access to water and provided with 1 flake of hay at 10PM. Blood collection vials intended

for GLP-1 measurement contained a DPP4 inhibitor[a] to prevent degradation of GLP-1. Samples were kept chilled throughout the sample collection period and processed immediately following completion of the sample collection period. Blood samples were centrifuged (1000 x g for 10 minutes) and plasma and serum were stored at -80°C. Hay, pasture, and grain samples were obtained for dietary analysis. Pasture samples were shipped overnight to our lab on ice. All diet samples were then stored at -20° or -80°C until being shipped to Equi-Analytical Laboratories (Ithaca, NY) for analysis. Pasture samples were shipped to the Equi-Analytical laboratory overnight on dry ice.

Body condition scores were assessed using the system developed by Henneke et al.[118] Age, breed, sex, and clinical group status (non-obese/no prior laminitis history, non-obese/prior laminitis history, obese/no prior laminitis history, and obese/prior laminitis history) were documented. Horses with a BCS ≥ 7 were classified obese.[118] Daily caloric consumption (Mcal) and grams crude protein (CP), neutral detergent fiber (NDF), starch, and water soluble carbohydrates (WSC) consumed per kg bodyweight (bwt) were calculated based on dry matter weight of each dietary component determined from analysis of dietary samples multiplied by the daily dry matter (DM) weight fed of each dietary component. Dry matter weight of daily pasture consumption was calculated by multiplying the determined pasture dry matter proportion by an estimate of the amount of pasture consumed daily based on time spent on pasture. Grams of dry matter intake per hour spent on pasture was extrapolated from data indicating horses on pasture for 3,6,9, and 24 hours consumed 1.96, 1.52, 1.12, and 0.57 g DM intake per kg bwt per hour.[126] A linear regression model was fit with log transformed DM intake values as an outcome and hours on pasture as a predictor, see Equation 2.1. Gram dry matter intake estimates where reduced by 75% for horses wearing grazing muzzles at time of sampling.[127] See Section 2.3.2 for description of GLU, INS, TG, NEFA, LEP, and APN assays. DPP4 and GLP-1 assays are described below.

DPP4 assay

The DPP4/CD26 Assay kit for Biological Samples [BML-AK498] from Enzo Life Sciences per manufacturer's instructions and a fluorogenic substrate (H-Gly-Pro-AMC) was used to determine DPP4 activity in the samples. Each sample was run in duplicates

[a]Millipore DPP4-010

and a Spectra Max Genini EM fluorescence plate reader with an excitation/emission wavelength of 380 nm/460nm was used to capture the changes in fluorescence over time. Activity is expressed as Relative Fluorescence Units (RFU) per minute and was calculated by determining the slope of the line of RFU versus time for each sample.

GLP-1 assays

GLP-1 is rapidly metabolized and inactivated by the enzyme DPP4. Thus, it is important to distinguish between measurements of the intact active hormone or the sum of the intact active hormone and its inactive metabolites.

The GLP-1 total ELISA [EZGLP1T-36K] from EMD Millipore Corporation per manufacturer's instructions was used. Each sample was run in duplicate. Absorbance was read at 450nm and 590 nm in a Spectra Max Plus plate reader and the concentration of GLP-1 total was determined using the SoftMax Pro 6.2.1 program with a 4-parameter logistic curve-fit.

The Glucagon-Like Peptide-1 (active) ELISA [EGLP-35K] from EMD Millipore Corporation per manufacturer's instructions was used to determine the concentration of GLP-1 active in the samples. Each sample was run in duplicate and a Spectra Max Genini EM fluorescence plate reader with an excitation/emission wavelength of 355 nm/460nm was used. SoftMax Pro 6.2.1 program was used to fit RFU to the concentration with a cubic-spline curve fit.

3.3.2 Variant discovery and genotyping

DPP4

The coding region of *DPP4* was sequenced from cDNA in 2 EMS suspect cases (muscle) and 1 horse (liver) of unknown status. Primers used to sequence *DPP4* from cDNA are listed in Table B.1. A single missense mutation identified in exon 11 located at chr18:42522956 (EquCab2 assembly, dbSNP rs397231534[b]). The SNP was genotyped in the entire sample population. Genomic DNA was extracted from whole blood using a commercially available kit.[c] NEBcutter[d] was used to identify differences in restriction enzyme sites between the reference sequence of *DPP4* and the single SNP. An enzyme *Cac81* was determined to cut the variant sequence. The equine reference sequence was used to design primers to amplify an 830 base pair product encompassing the variant (primers reported in Table B.1). Standard PCR amplification was performed with 35 cycles and a 58.5°C annealing temperature. The PCR product was incubated with 0.75 units of *Cac81*[e] enzyme at 37°C for 6 hours.

GCG

Standard Sanger sequencing of proglucagon (*GCG*) exons and 5' flanking region was performed in two horses with a history of endocrinopathic laminitis (primers reported in Table B.1). A single variant was identified in non-coding exon 1 (5'UTR). Sequencing of the region was repeated in 6 additional horses followed by development of an RFLP assay to genotype the entire sample population. Restriction enzyme *BanI* was determined to cut the variant sequence. The variant was genotyped in the entire sample population using the same PCR amplification protocol described above, the PCR product was incubated with 3 units of *BanI* enzyme at 37°C for 6 hours.

Standard Sanger sequencing of the 5' *GCG* flanking region containing the ERV LTR insertion revealed five variants, therefore Sanger sequencing of the region was elected over designing restriction fragment length polymorphism genotyping assays to genotype

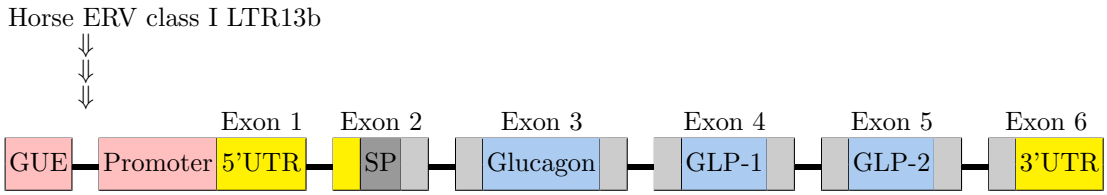
[b]Database of Single Nucleotide Polymorphisms (dbSNP). Bethesda (MD): National Center for Biotechnology Information, National Library of Medicine. http://www.ncbi.nlm.nih.gov/SNP/

[c]Puregene blood core kit, Qiagen Sciences, Germantown, MD.

[d]http://tools.neb.com/NEBcutter2/

[e]New England BioLabs, Ipswich, MA.

Figure 3.1: Proglucagon gene structure (ERV, endogenous retrovirus; LTR, long terminal repeat; GUE, proglucagon upsteam enhancer region; UTR, untranslated region; SP, signal peptide; GLP, glucagon-like peptide).



the entire population.

Haplotypes for the *GCG* variants were estimated using the implementation of the EM algorithm coded into the haplo.stats R package.[f]

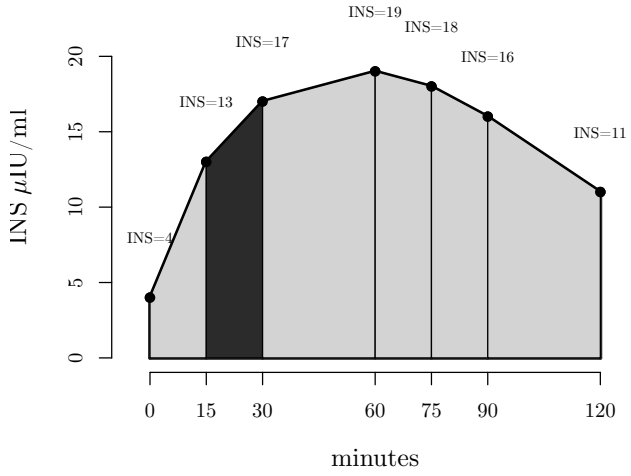
3.3.3 Statistical Methods

Single summary statistics:

Area under the curve (AUC) is a common method of integrating a response measured repeatedly over a period of time into a single summary statistic. AUC calculated by the trapezoidal method is frequently used to assess insulin and glucose dynamics during an oral glucose tolerance test.[173] The trapezoidal method sums the trapezoidal areas between measurement time points yielding a single summary statistic, the area under the curve (Figure 3.2). An integrated assessment often provides more utility than measurement of stimulated insulin secretion at a single time point or peak concentration. Although AUC is a simple, useful summary statistic it does not adequately describe the dynamic nature of the responses of interest in oral glucose stimulation. Deficiencies of AUC include it being largely influenced by the baseline measurement which can hinder interpretation of the dynamic response to a stimulus. Two experimental groups may be misinterpreted as having a significantly different response to a stimulus when in reality they only differed in their baseline values and had a very similar response to the stimulus. This issue can be circumvented by calculating a positive incremental AUC which subtracts the baseline measurement from all values when calculating the

[f]Sinnwell and Schaid, 2005, http://www.mayo.edu/research/documents/manualhaplostatspdf/doc-10026895.

Figure 3.2: Area under the curve trapezoidal rule calculation. The area under the curve is approximated by summing the trapezoidal areas between measurement time points.



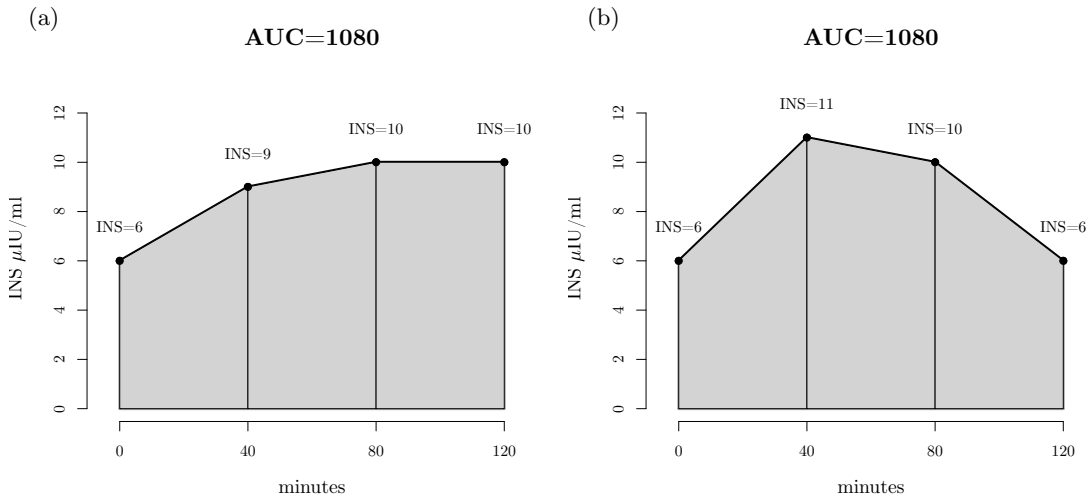
area under the curve. Secondly, it is possible for two different curves to have the same AUC resulting in a potential loss of important biological information. For example, the curve featured in Figure 3.3a exhibits a gradual initial increase in insulin without return to baseline values while the curve featured in Figure 3.3b demonstrates a rapid initial rate of change and rapid deceleration in order to return to baseline values. Interestingly these two differently shaped curves have identical area under the curves illustrating the information lost by assessing the area under the curve.

The effect of variables of interest on the area under the curve were tested one at a time using a multivariate, multilevel model in which a random intercept was included for each farm and the covariance of GLU, INS, GLP1a, and GLP1tot area under the curves was estimated at the farm and individual level. An overall test of the variable of interest's effect on the mean AUC was determined by performing a Wald chi-square test.

Trajectory analysis:

Trajectory analysis (a.k.a. longitudinal analysis, growth curve analysis) models the response curve as a function of time facilitating a more thorough examination of the

Figure 3.3: Demonstration of different curves with identical area under the curve (AUC)

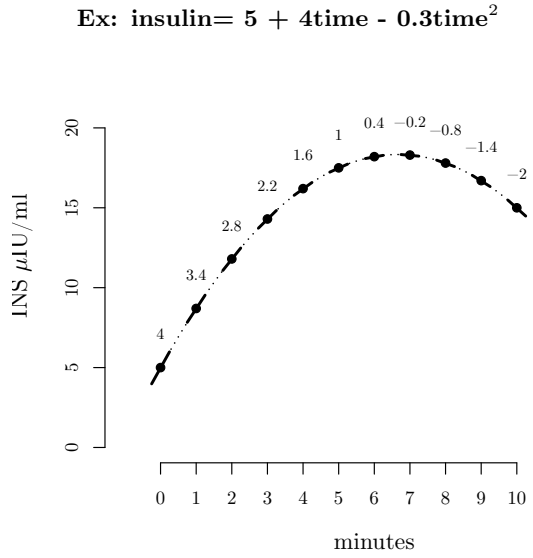


dynamic response to a glucose stimulus. The response variables of interest in an oral glucose test typically exhibit a "rise-fall" pattern (a quadratic relationship with time). Multilevel, longitudinal analysis facilitates a more thorough examination of the dynamic response to a glucose stimulus. Modeling the response curve as a quadratic function of time is well-suited for examining variation in baseline values, the initial rate of change, and deceleration of the rate of change. Multilevel longitudinal analyses account for the correlation of repeated measurements within individuals and correlation of observations of individuals sampled from the same farm.

Non-linearity in trajectories was checked by plotting and significance tests of quadratic and cubic fixed time effects. AIC and likelihood ratio tests were performed to determine inclusion of the appropriate individual and farm level random time effects. The most parsimonious model included fixed linear and quadratic time effects, a random intercept at the individual and farm level, and random linear and quadratic time effects at the individual levels. Covariance of GLU, INS, GLP1a, and GLP1tot intercepts were modeled at the farm level and covariance of GLU, INS, GLP1a, and GLP1tot intercepts, linear, and quadratic slopes were modeled at the individual level. Time was centered at time point zero-minutes.

The effect of variables of interest on the mean trajectory were tested one at a time by modeling their interaction effects with the intercept, linear time slope, and quadratic

Figure 3.4: Quadratic time function example. When time is centered at zero for a quadratic function of time ($y=B_0 + B_1 time + B_2 time^2$), B_0=y value at time zero, B_1=rate of change at time zero, B_2=one half of the acceleration or deceleration rate (degree to which B_1 is reduced at each time point)



time point	deceleration rate($2*B_2$)	old rate - ($2*B_2$)	new rate
1	0.3*2=0.6	4-0.6	3.4
2	0.3*2=0.6	3.4-0.6	2.8
3	0.3*2=0.6	2.8-0.6	2.2
4	0.3*2=0.6	2.2-0.6	1.6
5	0.3*2=0.6	1.6-0.6	1
6	0.3*2=0.6	1-0.6	0.4
7	0.3*2=0.6	0.4-0.6	-0.2
8	0.3*2=0.6	-0.2-0.6	-0.8
9	0.3*2=0.6	-0.8-0.6	-1.4
10	0.3*2=0.6	-1.4-0.6	-2

time slope. See Section 1.3 for an overview of multivariate, longitudinal, multilevel models. An overall test of the variable of interest's effect on the mean trajectory was determined by performing a Wald chi-square test. Accordingly, p values relate to curve differences in intercept, linear slope, and/or quadratic slope. Estimated marginal means predicted from the models were used for graphical representation of the trajectories. Statistical analyses were performed in R (version 15.3)[g] using ASReml-R[129] statistical software.

[g]R Core Team (2012). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, http://www.R-project.org/.

3.4 Results

3.4.1 Descriptive statistics

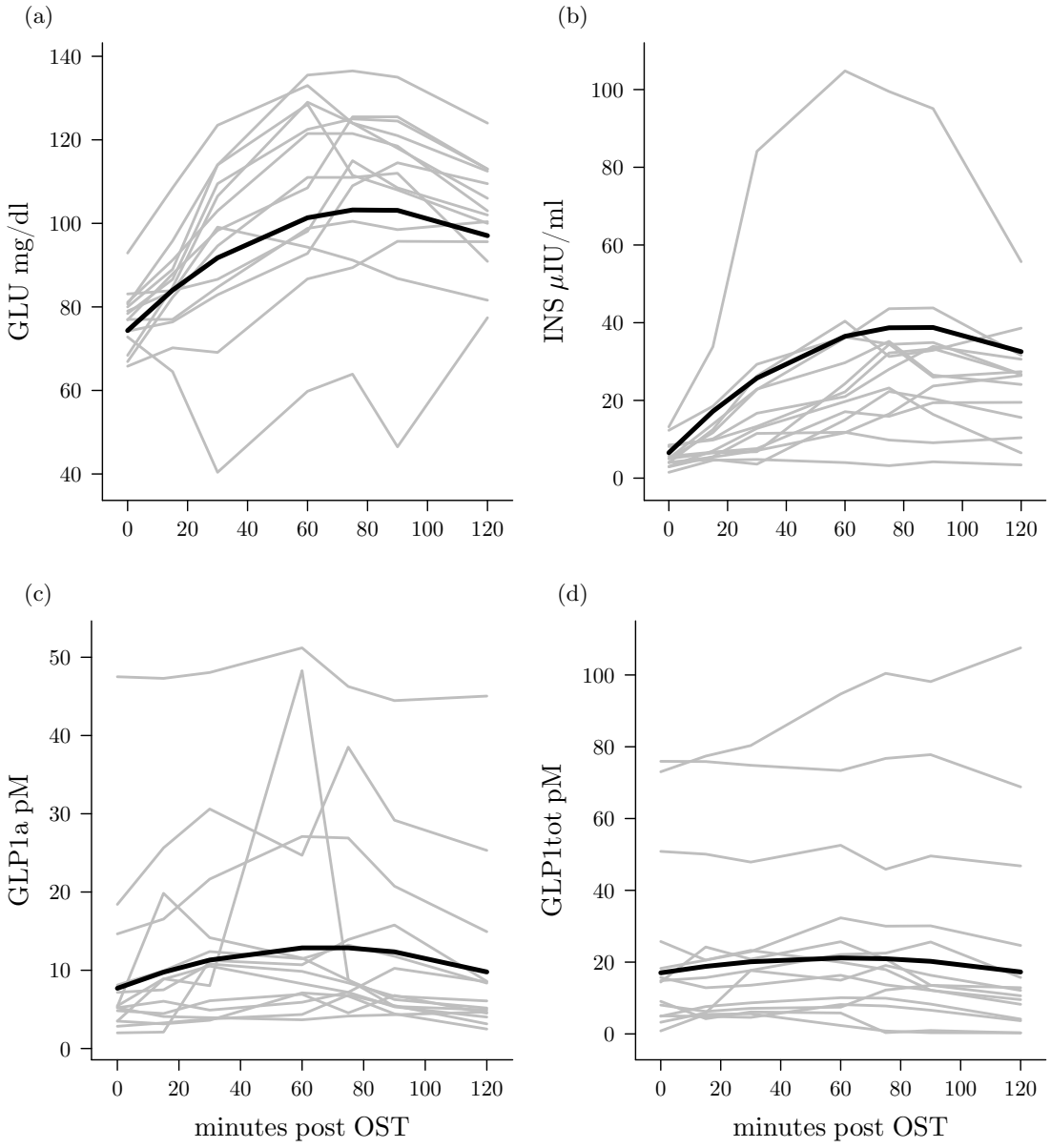
Characteristics of the sampled farm populations are presented in Table 3.0. Figure 3.5 illustrates the inter-individual variability in the glucose, insulin, and GLP-1 responses to an oral sugar challenge. It is readily apparent that individual trajectories are often not similar to the population mean trajectory indicated by the bold black line. Inter-individual variability was observed about the intercept (measurement value at time zero just prior to oral sugar administration), the initial rate of change, and deceleration of the initial rate of change (determines time till baseline return). The plotted data, in addition to AIC values, supported the use of random quadratic time effect model.

Table 3.0: Descriptive statistics of the sampled farms

	MorganFarm1	MorganFarm2	MorganFarm3	MorganFarm4	WelshFarm1	WelshFarm2	WelshFarm3
N=	16	20	8	12	12	19	25
Age: mean(sd)	12.1(6.3)	10.2(3.8)	17.8(5.2)	11.9(5.2)	10.2(3.4)	14.4(8.4)	12.4(6.3)
Gender:							
gelding/mare/stallion	5/10/1	9/10/1	0/7/1	0/12/0	2/10/0	3/14/2	3/22/0
Clinical group:							
$O^-L^-/O^-L^+/O^+L^-/O^+L^+$	10/1/4/1	9/4/5/2	7/1/0/0	5/2/5/0	5/1/4/2	17/2/0/0	0/11/0/14
Diet mean(sd)							
Mcal /kg bwt	0.05(0.01)	0.03(0.01)	0.06(0.01)	0.04(0.00)	0.03(0.00)	0.03(0.00)	0.03(0.00)
CP g/kg bwt	2.24(0.66)	1.30(0.40)	2.86(0.25)	3.82(0.10)	1.90(0.00)	2.33(0.19)	3.70(0.30)
NDF g/kg bwt	15.48(2.26)	7.23(1.53)	15.10(2.16)	9.48(0.30)	6.90(0.00)	6.36(0.24)	6.50(0.31)
WSC g/kg bwt	3.84(0.62)	1.41(0.58)	1.32(0.28)	2.02(0.05)	1.64(0.00)	0.86(0.35)	1.05(0.11)
Starch g/kg bwt	0.60(0.17)	1.37(0.32)	0.34(0.15)	0.08(0.01)	0.50(0.00)	0.48(0.01)	0.20(0.02)
Biochemical mean(sd)							
TG mg/dl	25.57(10.56)	18.18(5.50)	32.07(9.71)	40.27(31.93)	58.17(33.04)	20.53(7.49)	114.53(89.06)
NEFA mmol/L	0.24(0.18)	0.24(0.11)	0.22(0.09)	0.41(0.17)	0.52(0.14)	0.22(0.11)	0.49(0.13)
LEP ng/ml	6.03(4.23)	3.99(3.68)	7.10(3.68)	5.20(3.88)	3.40(2.43)	7.08(5.02)	5.18(2.28)
APN ng/ml	5557(2626)	7257(1858)	3884(1527)	2955(2633)	6152(4508)	6789(3906)	3770(3110)
DPP4 activity	12.79(3.69)	10.67(2.94)	24.3(3.51)	19.66(3.66)	35.26(13.15)	21.86(11.8)	25.2(7.17)

O^-L^-=non-obese/no prior laminitis, O^-L^+=non-obese/prior laminitis, O^+L^-=obese/no prior laminitis, O^+L^+=obese/prior laminitis, Mcal=megacalories, kg=kilogram, bwt=bodyweight, NDF=neutral detergent fiber, WSC=water soluble carbohydrates, TG=triglycerides, NEFA=non-esterfied fatty acids, LEP=leptin, APN=adiponectin, DPP4= dipeptidyl peptidase-4

Figure 3.5: Population mean (black) and random sample (grey) of 15 raw trajectories for glucose (a), insulin (b), active GLP-1 (c), and total GLP-1 (d) during an oral sugar test (OST). Oral sugar administered at time point zero.



3.4.2 *DPP4* and *GCG* variants

Table 3.1 reports the location and frequencies of the genotyped variants. A single missense (Thr>Ala) mutation was identified in exon 11 of the *DPP4* gene located at chr18:42522956 (EquCab2 assembly, dbSNP rs397231534[h]). Minor allele (G) frequencies in the entire, Welsh pony, and Morgan horse populations were 0.45, 0.54, and 0.36, respectively.

For the *GCG* gene, a SNP was identified in non-coding exon 1 (5'UTR) in addition to five variants identified in the 5' flanking region containing the ERV LTR insertion. Haplotypes for the *GCG* variants were estimated and will be referred to as haplotypes A-E. Haplotypes A-D had frequencies greater than 5% (*haplotype A*:CCAGGC, frequency=0.31; *haplotype B*:CCAG**C**C, frequency=0.27; *haplotype C*:C**A**A**T**G**G**, frequency=0.27; *haplotype D*:**G**CAGGC, frequency=0.06, *haplotype E* denotes ten pooled rare haplotypes (frequency<5%).

Table 3.1: Description of *DPP4* and *GCG* genotyped variants

					Minor Allele Frequency		
						Breed	
Gene	Location	chr:bp	Allele change	Residue change	Entire population	Morgan	Welsh Pony
DPP4	exon 11	18:42386473	A>G	Thr>Ala	0.45	0.36	0.53
GCG	GUE	18:42523969	C>G	NA	0.09	0.15	0.02
GCG	GUE	18:42523797	C>A	NA	0.30	0.23	0.21
GCG	GUE	18:42523699	A>G	NA	0.04	0.06	0.05
GCG	GUE	18:42523690	G>T	NA	0.31	0.23	0.23
GCG	GUE	18:42523627	G>C	NA	0.33	0.50	0.21
GCG	5'UTR	18:42522905	C>G	NA	0.33	0.27	0.22

chr, chromosome; bp, base pair; DPP4, dipeptidyl-peptidase 4 gene; GCG, proglucagon gene; GUE, proglucagon gene upstream enhancer; UTR, untranslated region

[h]Database of Single Nucleotide Polymorphisms (dbSNP). Bethesda (MD): National Center for Biotechnology Information, National Library of Medicine. http://www.ncbi.nlm.nih.gov/SNP/

3.4.3 DPP4 activity

DPP4 activity (RFU/min) was tested for association with the following univariate predictors: age, gender, clinical group, diet , *DPP4* genotype, and blood concentration of triglyceride, NEFA, leptin, and adiponectin (Table 3.2). Age and adiponectin were negatively correlated with DPP4 activity ($p < 0.05$, $p < 0.01$; respectively) and DPP4 activity was significantly higher in Welsh ponies compared to Morgan horses (+10.66(±4.93) units, $p < 0.05$).

Table 3.2: DPP4 activity (RFU/min) fixed effect estimate(se)

univariate predictor	estimate(se)
Ref group (mean age)	21.41(3.16)
1 SD increase in age	-1.64(0.76)*
Ref group (gelding)	20.30(3.48)
mare	1.28(2.00)
stallion	0.73(3.95)
Ref group (Morgan breed)	16.65(3.26)
Welsh Pony	10.66(4.93)*
Ref group (non-obese, no prior laminitis)	22.15(3.26)
non-obese, prior laminitis	-2.70(2.38)
obese, no prior laminitis	-1.87(2.30)
obese, prior laminitis	-0.21(2.84)
Ref group(mean Mcal)	21.42(3.15)
1 SD increase in Mcal	-0.42(1.43)
Ref group (mean nutrients)	21.09(4.11)
1 SD increase in Starch	-2.17(2.34)
1 SD increase in CP	-2.32(2.33)
1 SD increase in NDF	-0.65(3.63)
1 SD increase in WSC	3.69(3.33)
Ref group (mean triglycerides)	21.29(3.20)
1 SD increase in triglycerides	-0.65(0.95)
Ref group (mean NEFA)	21.34(3.05)
1 SD increase in NEFA	0.44(0.99)
Ref group (mean leptin)	21.33(3.13)
1 SD increase in leptin	0.31(0.77)

Table 3.2 *Continued on next page*

Table 3.2 *Continued from previous page*

univariate predictor	estimate(se)
Ref group (mean adiponectin)	21.25(3.06)
1 SD increase in adiponectin	-2.30(0.79)**
Ref group (*DPP4* 0 copies)	20.41(3.18)
DPP4 1 copy minor allele	1.02(1.06)
Ref group (*GCG* haplotype A)	22.49(3.62)
GCG haplotype B	-0.67(1.57)
GCG haplotype C	-1.76(1.46)
GCG haplotype D	0.69(2.72)
GCG haplotype E	0.09(2.06)

Asterisks indicate significance of beta estimates (*<0.05, **<0.01, ***<0.001). *GCG* haplotype A=highest frequency haplotype, *GCG* haplotype E=pooled rare haplotypes with frequency < 0.05

3.4.4 Glucose, insulin, and GLP-1 area under the curves and trajectories

Multiple factors were tested for univariate association with glucose (Table 3.3), insulin (Table 3.4), active GLP-1 (Table 3.5) and total GLP-1 (Table 3.6) area under the curves and trajectories during an oral sugar test. In numerous instances, trajectory analysis detected associated factors that went undetected by an area under the curve analytical approach, hence only the trajectory analysis results are described below. Estimated effects of the different factors on glucose, insulin, active GLP-1, and total GLP-1 AUCs and trajectory intercepts, linear slopes (unit change per 15 minute interval), and quadratic slopes (1/2 the deceleration rate) from the univariate predictor model are presented in Tables B.2-B.5. Factors with a p-value significance level <0.2 in the univariate predictor analysis were included in a multiple predictor regression analysis of glucose, insulin, and GLP-1 trajectories (Appendix Tables B.6 to B.9).

Glucose

Factors demonstrating a significant association with the glucose trajectory included: age, breed, clinical group, caloric intake, nutrient intake, and NEFA blood concentration

(Tables 3.3, B.2). Older individuals had a lower initial rate of increase in glucose and a corresponding reduction in deceleration of the glucose rate (Figure 3.6a). Welsh ponies had a higher initial rate of increase in glucose and a corresponding increase in deceleration of the glucose rate in comparison to the Morgan horses (Figure 3.6b). Glucose trajectories differed by clinical group status. In comparison to the reference group (non-obese horses with no prior history of laminitis), obese horses with no prior laminitis history had a higher baseline value of glucose, although their response following oral sugar administration was not significantly different. Obese horses with a prior history of laminitis did not differ at the baseline glucose level but did have a higher initial rate of increase in glucose and a corresponding increase in deceleration of the glucose rate (Figure 3.6c). The initial glucose rate of increase was positively correlated with caloric intake per kg bodyweight (Figure 3.7a). Diet composition was not associated with differences in baseline glucose levels (Figure 3.7b), however higher fiber diets had a reduced glycemic response whereas diets high in water soluble carbohydrates had an elevated glycemic response. A high starch diet was associated with a mild reduction in the glycemic response. Elevated blood concentration of triglycerides and NEFA were associated with an increased glycemic response (Figure 3.8a,b).

Insulin

The insulin trajectory was significantly associated with breed, clinical group, nutrient intake, triglycerides, leptin, adiponectin, and DPP4 activity (Tables 3.4, B.3). Baseline values of insulin were not significantly different by breed, however the insulinemic response was much higher in Welsh ponies than in Morgan horses (Figure 3.9a).
Insulin trajectories varied with obesity and prior laminitis status (Figure 3.9b). Obese horses with a prior history of laminitis had elevated baseline insulin values and a higher insulinemic response. The insulinemic response was also elevated in non-obese horses with a prior history of laminitis, although the difference from non-obese horses with no prior history of laminitis did not reach statistical significance. Horses with higher starch or higher fiber in their diets tended to have a reduced insulinemic response (Figure 3.9c). Triglyceride, leptin, and DPP4 activity blood levels were positively correlated with insulinemic responses whereas adiponectin blood levels were negatively correlated with the insulinemic response (Figures 3.10a-d).

Figure 3.6: Predicted trajectories of glucose during an oral sugar test (OST) by age (a), breed (b), and clinical group (c). Oral sugar administered at time point zero. Solid lines are estimated trajectories and dashed lines represent standard error of the mean.

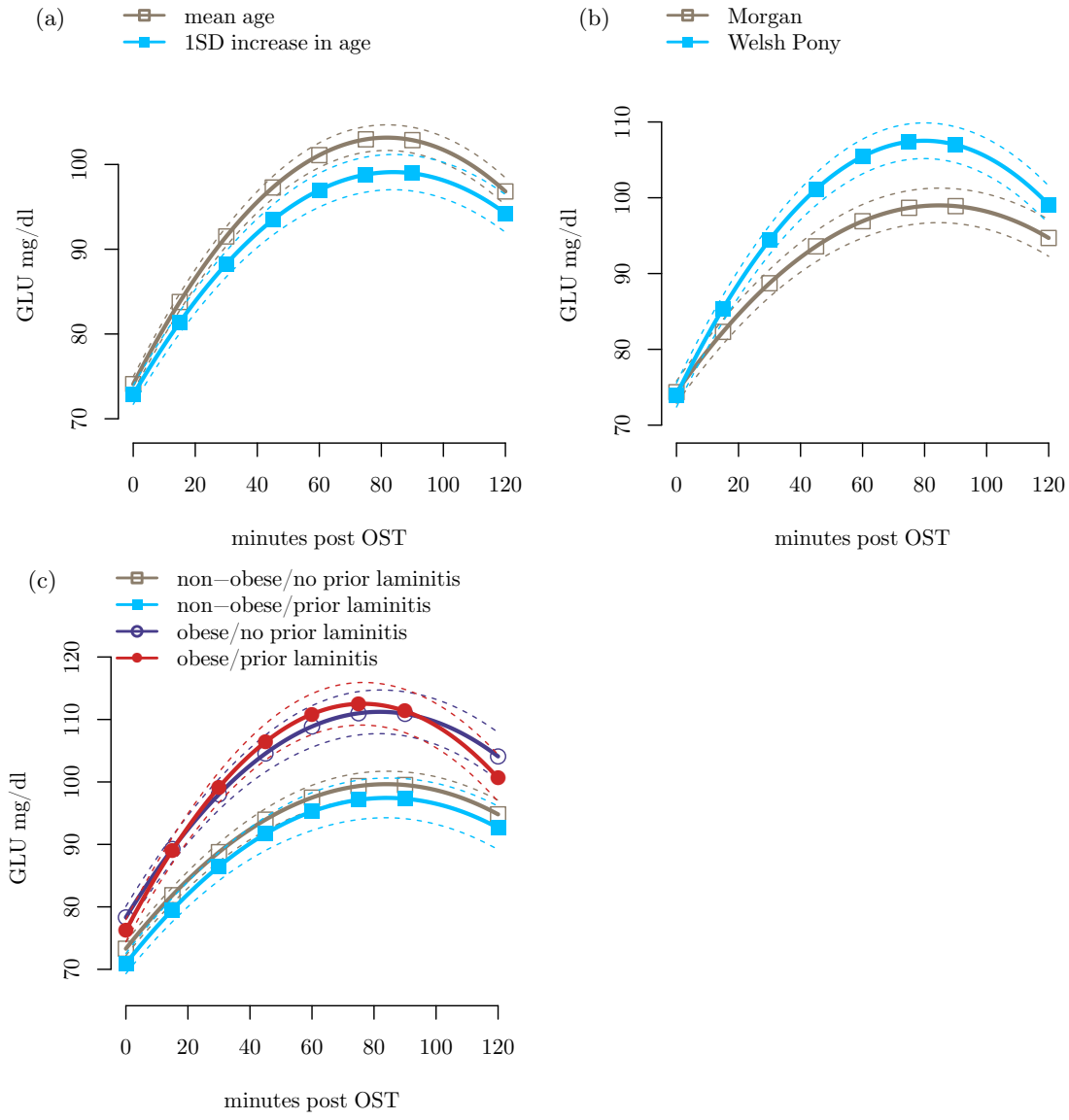


Table 3.3: Comparison of glucose AUC and trajectory association tests with univariate predictors of glucose response to an oral sugar test

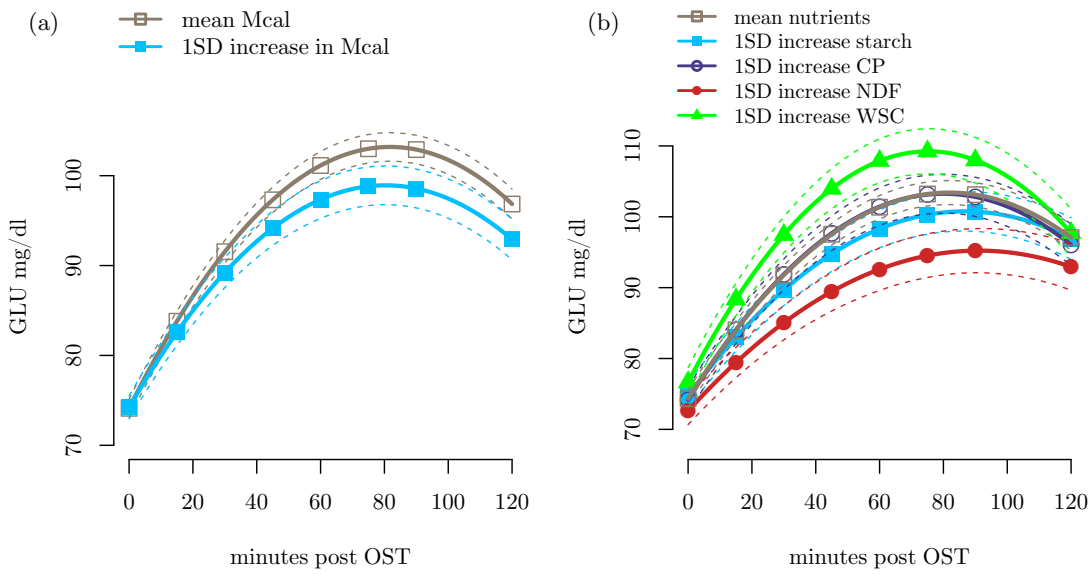
univariate predictor	glucose AUC (wald χ^2, df) p-value	glucose trajectory (wald χ^2, df) p-value
age	**(6.7, 1df)=9.4E-03**	**(9.6, 3df)=2.2E-02**
sex	(0.2, 2df)=9.1E-01	(3.7, 6df)=7.2E-01
breed	(2.2, 1df)=1.4E-01	**(12.8, 3df)=5.2E-03**
clinical group	**(17.9, 3df)=4.6E-04**	**(28.7, 9df)=7.4E-04**
caloric intake	**(4.1, 1df)=4.3E-02**	**(11.3, 3df)=1.0E-02**
nutrient intake	(4.8, 4df)=3.1E-01	**(30.9, 12df)=2.0E-03**
triglycerides	(0.4, 1df)=5.4E-01	**(8.3, 3df)=4.0E-02**
NEFA	**(6.6, 1df)=1.0E-02**	**(18.6, 3df)=3.2E-04**
leptin	(3.3, 1df)=6.9E-02	(1.5, 3df)=6.9E-01
adiponectin	(0.0, 1df)=8.4E-01	(2.2, 3df)=5.4E-01
DPP4 activity	(1.5, 1df)=2.3E-01	(7.6, 3df)=5.5E-02
DPP4 genotype	(0.0, 1df)=8.6E-01	(2.2, 3df)=5.2E-01
GCG haplotype	(3.3, 4df)=5.0E-01	(20.1, 12df)=6.5E-02

Results obtained from a multivariate response model (responses included glucose, insulin, active GLP-1 and total GLP-1 measured at 7 time points over a 2 hour period)

GLP-1

Factors associated with GLP-1 trajectories included sex, clinical group, diet composition, NEFA blood levels, and haplotype at the proglucagon gene (*GCG*) upstream enhancer region (GUE) (Tables 3.5, 3.6, B.4, B.5). Active and total GLP-1 unexpectedly decreased throughout the duration of the oral sugar tests in stallions, although baseline GLP-1 levels tended to be higher (Figure 3.11a). Higher starch and higher protein diets were associated with a reduced GLP-1 response, although individuals with higher starch diets tended to have higher basal GLP-1 levels and individuals with higher protein diets tended to have lower basal GLP-1 levels (Figure 3.11b). Basal GLP-1 levels were significantly lower in non-obese horses with a prior history of laminitis and also trended toward being lower in obese horses with and without a history of laminitis. However, obese horses with no prior laminitis had a significantly greater positive GLP-1 response following oral sugar administration whereas obese horses with a history of laminitis had decreasing levels of total GLP-1 throughout the oral sugar test. *GCG* haplotype was modestly associated with total GLP-1 trajectory variation (Figure

Figure 3.7: Predicted trajectories of glucose during an oral sugar test (OST) by caloric intake (a) and nutrient composition (b) per kg bodyweight. Oral sugar administered at time point zero. Solid lines are estimated trajectories and dashed lines represent standard error of the mean.

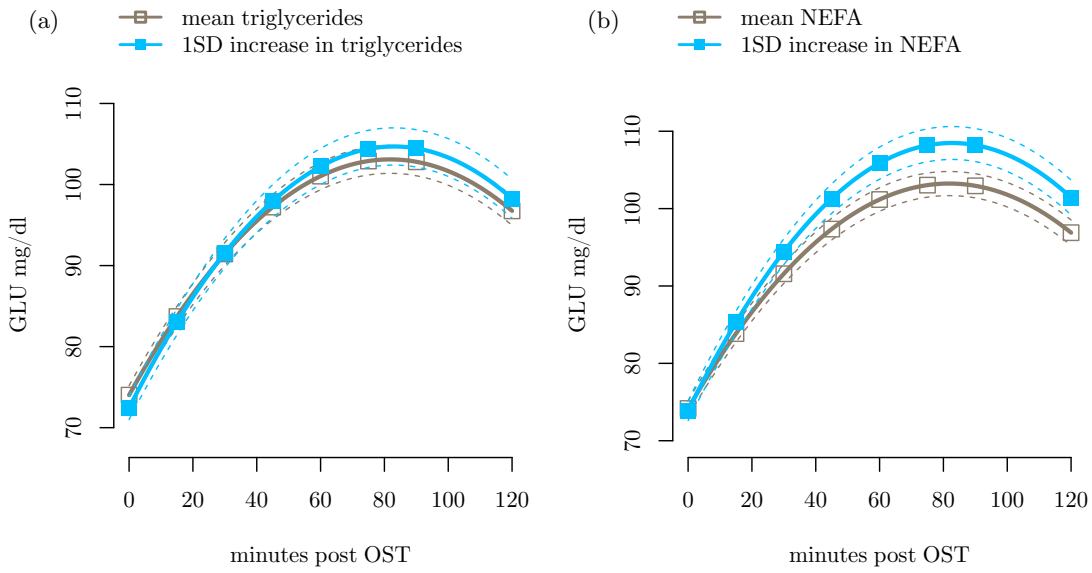


3.12a). The most common haplotype (referred to as "haplotype A" in Figure 3.12a) had the lowest basal total GLP-1 level. "Haplotype E" consisting of a pooling of rare haplotypes with a frequency less than 5% had a significantly higher basal total GLP-1 level. "Haplotype B" demonstrated decreasing total GLP-1 levels throughout the oral sugar test similar to obese horses with a history of laminitis.

Correlation of glucose, insulin, and GLP-1 trajectories

Estimate of trajectory variance and covariance from an unconditional multivariate longitudinal model were used to obtain estimates of correlation. The correlation estimates of glucose, insulin, and GLP-1 baseline values, initial rate of increase, and deceleration are presented in Figure 3.13. As expected, given the same overall "rise-fall" pattern observed for the different trajectories, the initial rate of increase and deceleration rate were strongly negatively correlated. Individuals with a higher initial rate of increase experienced faster deceleration in attempt to achieve a return to baseline values in a timely manner. GLP-1 active and total baseline values, initial rates of change, and

Figure 3.8: Predicted trajectories of glucose during an oral sugar test (OST) by triglyceride (a) and NEFA (b) concentration. Oral sugar administered at time point zero. Solid lines are estimated trajectories and dashed lines represent standard error of the mean.



deceleration rates were also strongly correlated indicated minimal variation in GLP-1 active total ratios throughout the oral sugar test. Insulin and glucose initial rates of change and deceleration were correlated however appreciably correlation was not evident between insulin and glucose with GLP-1 active and total trajectories. Active GLP-1 baseline values were correlated with both active and total initial rates of change (positive correlation) and deceleration (negative correlation). Likewise, total GLP-1 baseline values were correlated with both active and total initial rate of change and deceleration.

Figure 3.9: Predicted trajectories of insulin during an oral sugar test (OST) by breed (a), clinical group (b), and nutrient composition (c). Oral sugar administered at time point zero. Solid lines are estimated trajectories and dashed lines represent standard error of the mean.

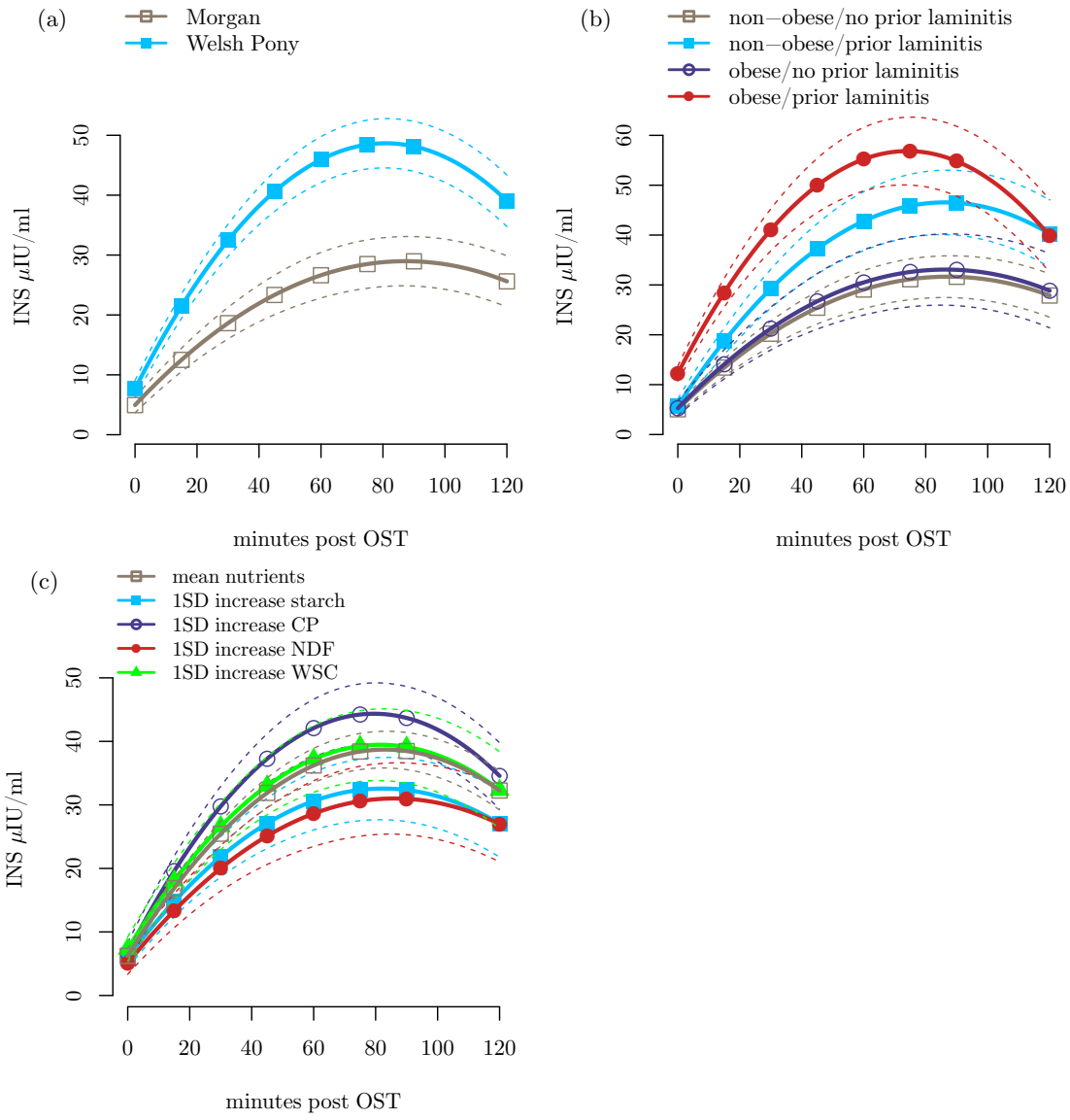


Figure 3.10: Predicted trajectories of insulin during an oral sugar test (OST) by triglyceride (a), leptin (b), adiponectin (c), and DPP4 activity (d) blood levels. Oral sugar administered at time point zero. Solid lines are estimated trajectories and dashed lines represent standard error of the mean.

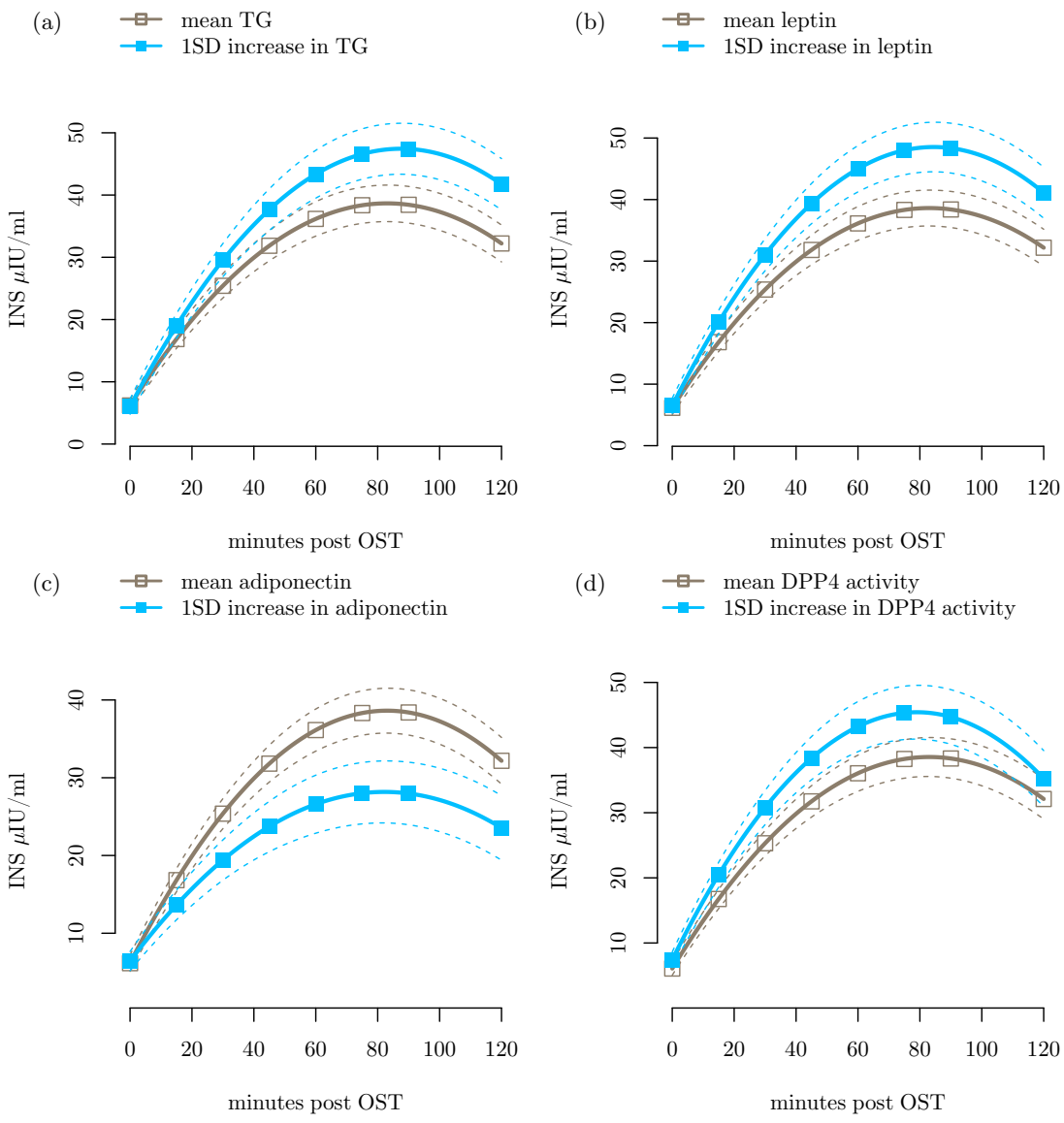


Figure 3.11: Predicted trajectories of active (a, c) and total (b, d) GLP-1 during an oral sugar test (OST) by sex and nutrient composition. Oral sugar administered at time point zero. Solid lines are estimated trajectories and dashed lines represent standard error of the mean.

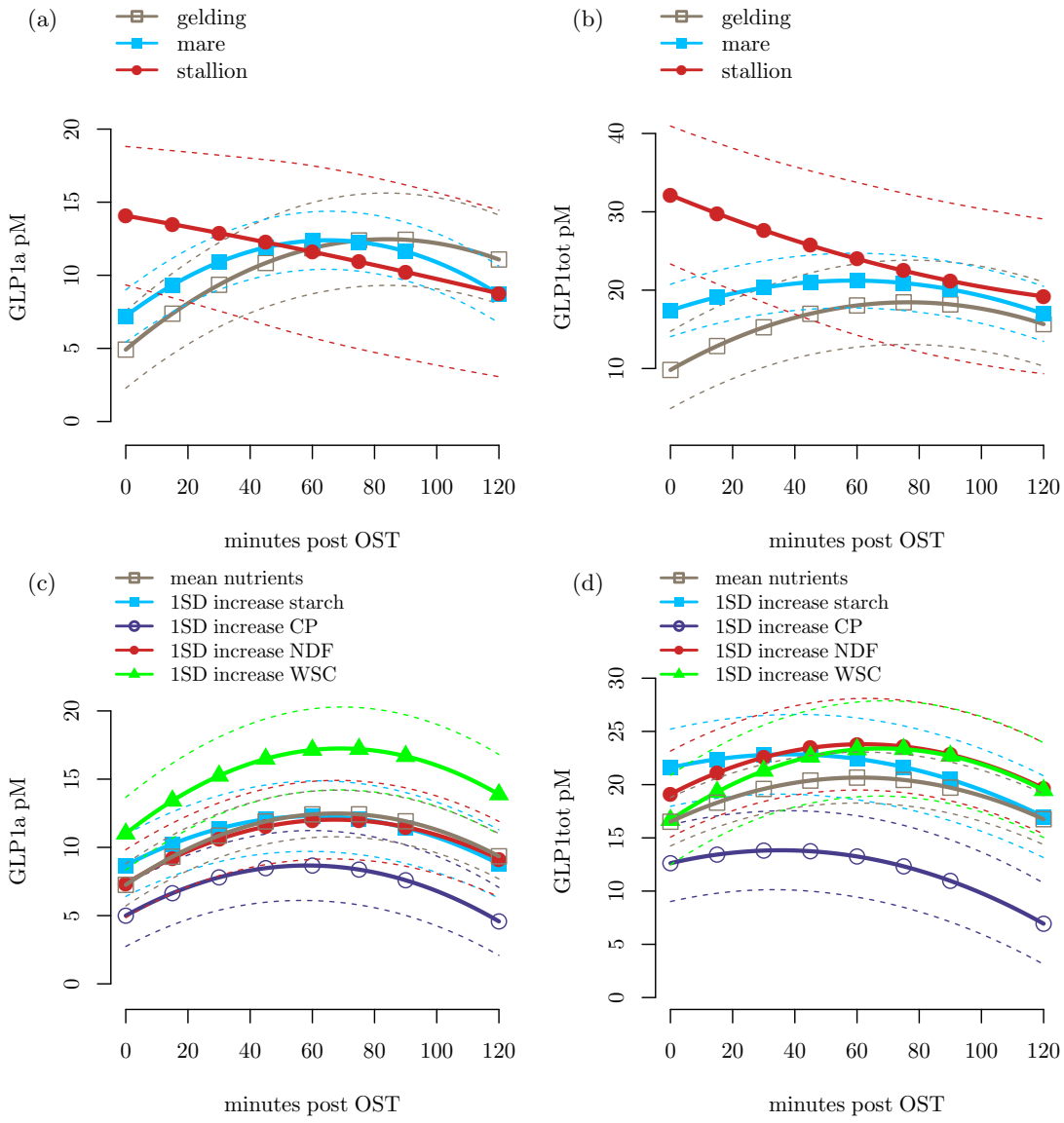


Table 3.4: Comparison of insulin AUC and trajectory association tests with univariate predictors of insulin response to an oral sugar test

univariate predictor	insulin AUC (wald χ^2, df) p-value	insulin trajectory (wald χ^2, df) p-value
age	(0.0, 1df)=9.7E-01	(4.0, 3df)=2.7E-01
sex	(1.6, 2df)=4.4E-01	(6.4, 6df)=3.8E-01
breed	(2.3, 1df)=1.3E-01	**(12.8, 3df)=5.2E-03**
clinical group	(7.2, 3df)=6.5E-02	**(28.7, 9df)=7.4E-04**
caloric intake	(0.8, 1df)=3.8E-01	(3.0, 3df)=3.9E-01
nutrient intake	**(26.3, 4df)=2.7E-05**	**(24.5, 12df)=1.7E-02**
triglycerides	(2.8, 1df)=9.6E-02	**(11.9, 3df)=7.8E-03**
NEFA	(0.0, 1df)=9.5E-01	(5.4, 3df)=1.4E-01
leptin	**(14.2, 1df)=1.6E-04**	**(13.0, 3df)=4.6E-03**
adiponectin	**(8.9, 1df)=2.9E-03**	**(15.1, 3df)=1.7E-03**
DPP4 activity	(1.1, 1df)=2.9E-01	**(10.1, 3df)=1.7E-02**
DPP4 genotype	(0.2, 1df)=6.3E-01	(6.5, 3df)=8.8E-02
GCG haplotype	(1.0, 4df)=9.2E-01	(11.5, 12df)=4.8E-01

Results obtained from a multivariate response model (responses included glucose, insulin, active GLP-1 and total GLP-1 measured at 7 time points over a 2 hour period)

Table 3.5: Comparison of active GLP-1 AUC and trajectory association tests with univariate predictors of active GLP-1 response to an oral sugar test

univariate predictor	active GLP-1 AUC (wald χ^2, df) p-value	active GLP-1 trajectory (wald χ^2, df) p-value
age	(0.1, 1df)=7.6E-01	(0.8, 3df)=8.4E-01
sex	(0.1, 2df)=9.7E-01	**(18.1, 6df)=6.0E-03**
breed	(0.7, 1df)=3.9E-01	(1.7, 3df)=6.5E-01
clinical group	(3.5, 3df)=3.3E-01	(15.9, 9df)=6.9E-02
caloric intake	(0.6, 1df)=4.3E-01	(1.7, 3df)=6.5E-01
nutrient intake	**(17.8, 4df)=1.4E-03**	**(21.7, 12df)=4.1E-02**
triglycerides	(0.1, 1df)=7.1E-01	(3.7, 3df)=3.0E-01
NEFA	(2.6, 1df)=1.1E-01	(3.7, 3df)=2.9E-01
leptin	(0.2, 1df)=7.0E-01	(1.0, 3df)=8.0E-01
adiponectin	(2.1, 1df)=1.5E-01	(6.2, 3df)=1.0E-01
DPP4 activity	(1.6, 1df)=2.1E-01	(2.3, 3df)=5.2E-01
DPP4 genotype	(0.0, 1df)=9.9E-01	(0.9, 3df)=8.2E-01
GCG haplotype	(2.2, 4df)=7.0E-01	(12.6, 12df)=4.0E-01

Results obtained from a multivariate response model (responses included glucose, insulin, active GLP-1 and total GLP-1 measured at 7 time points over a 2 hour period)

Table 3.6: Comparison of total GLP-1 AUC and trajectory association tests with univariate predictors of total GLP-1 response to an oral sugar test

univariate predictor	total GLP-1 AUC (wald χ^2, df) p-value	total GLP-1 trajectory (wald χ^2, df) p-value
age	(0.2, 1df)=7.0E-01	(2.4, 3df)=4.9E-01
sex	(0.7, 2df)=6.9E-01	**(22.4, 6df)=1.0E-03**
breed	(1.2, 1df)=2.8E-01	(4.0, 3df)=2.7E-01
clinical group	(4.7, 3df)=1.9E-01	**(25.3, 9df)=2.6E-03**
caloric intake	(0.1, 1df)=7.7E-01	(7.0, 3df)=7.1E-02
nutrient intake	**(22.0, 4df)=2.0E-04**	**(48.5, 12df)=2.5E-06**
triglycerides	(0.8, 1df)=3.8E-01	(6.5, 3df)=8.9E-02
NEFA	(3.1, 1df)=7.9E-02	**(8.1, 3df)=4.4E-02**
leptin	(0.0, 1df)=8.8E-01	(4.0, 3df)=2.7E-01
adiponectin	(1.9, 1df)=1.7E-01	(3.9, 3df)=2.7E-01
DPP4 activity	(0.7, 1df)=4.0E-01	(4.1, 3df)=2.5E-01
DPP4 genotype	(0.4, 1df)=5.2E-01	(0.8, 3df)=8.5E-01
GCG haplotype	(1.4, 4df)=8.5E-01	**(23.3, 12df)=2.5E-02**

*Results obtained from a multivariate response model (responses included glucose, insulin, active GLP-1 and total GLP-1 measured at 7 time points over a 2 hour period)

Figure 3.12: Predicted trajectories of total GLP-1 during an oral sugar test (OST) by clinical group (a) and *GCG* haplotypes (b). Oral sugar administered at time point zero. Solid lines are estimated trajectories and dashed lines represent standard error of the mean.

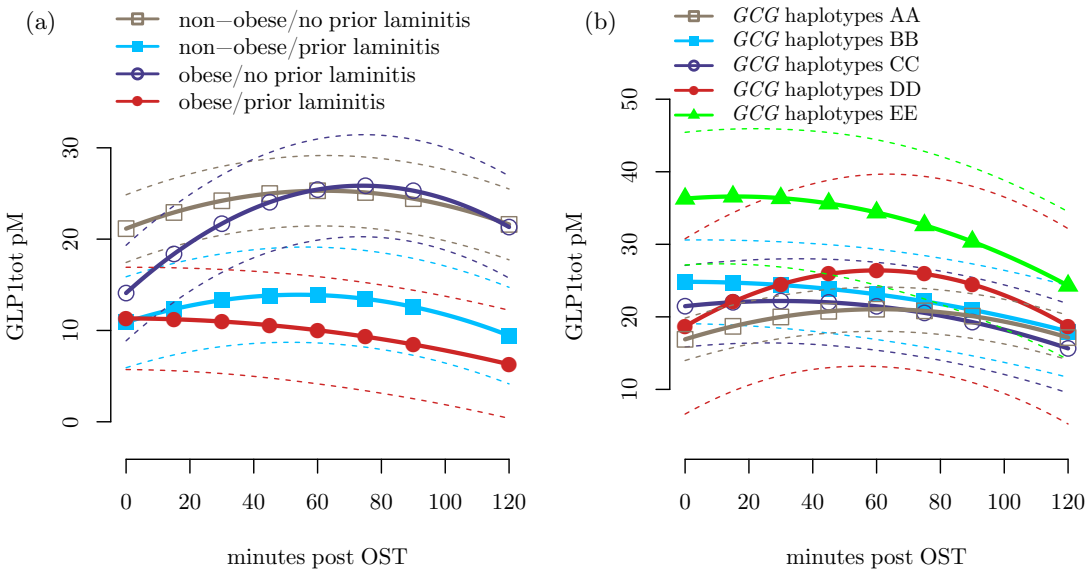
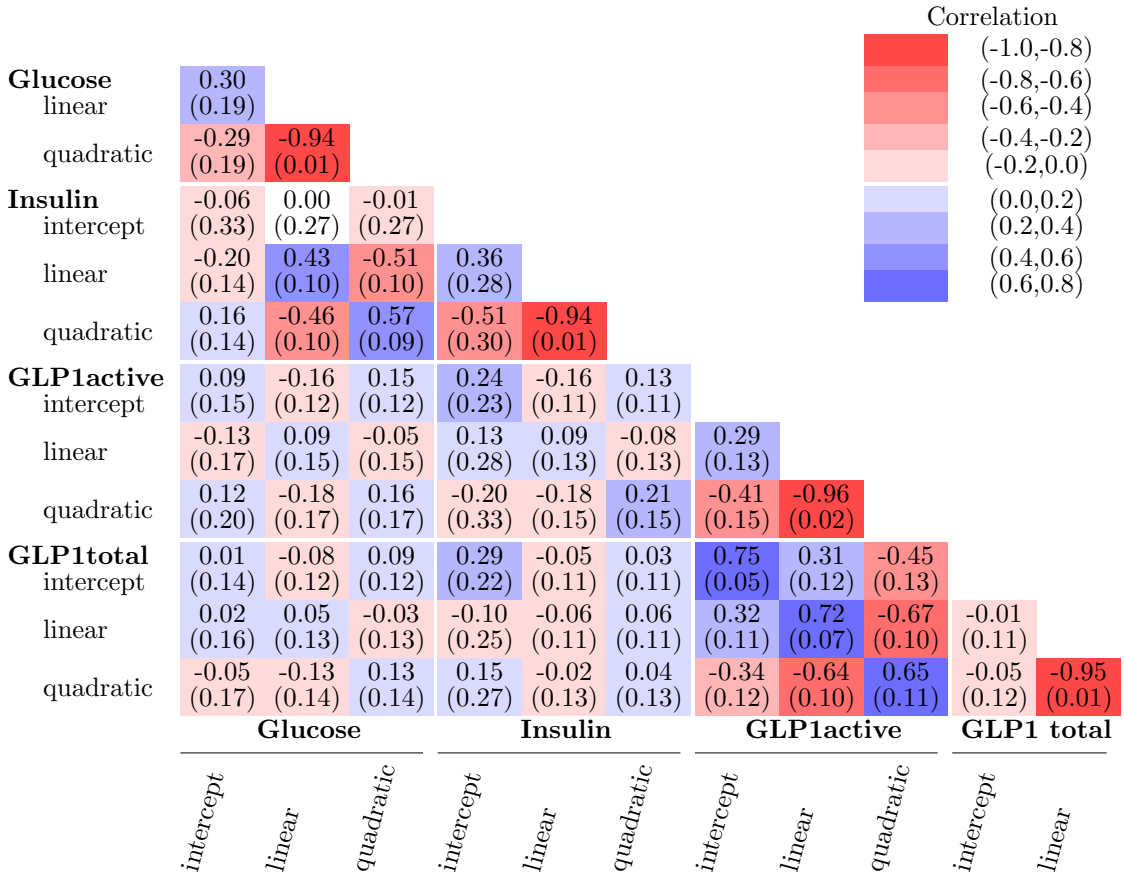


Figure 3.13: Correlation estimates (SE) for glucose, insulin, and GLP-1 trajectories. Correlation estimates determined from variance and covariance estimates obtained from an unconditional longitudinal model that included fixed and random linear and quadratic time effects and a random farm intercept. Time centered at 0 minutes. Standard errors estimated by the delta method.



3.5 Discussion

We hypothesized alteration in the incretin response may play a role in the pathogenesis of EMS. In this characterization study of the equine incretin response, a number of factors were found to be associated with the incretin hormone GLP-1 and insulin/glucose dynamics during the equine response to an oral sugar challenge. In addition, factors were also observed to be associated with DPP4, the major protease that breaks down GLP-1. It is important to note that tested factors reported in the main text of the chapter were included in both the DPP4 activity response model and the glucose/insulin/GLP-1 multivariate response model one at a time due to colinearity, other covariates were not included in the models. It was beyond the scope of this chapter to build an extensive prediction model. For the interested reader, estimates from a multiple predictor regression model are reported in Appendix B but should be interpreted with caution given the existence of colinearity among the tested predictors. Likewise, estimates reported in the main text can only be interpreted as correlated factors, not causally related factors. The objective of this study was to lay the groundwork in identifying factors associated with variation in the equine response to an oral sugar challenge which can later be built upon when attempting to determine causal pathways.

3.5.1 DPP4

Age and adiponectin levels were found to be negatively correlated with DPP4 blood activity levels. Similarly, DPP4 activity has been reported to be negatively correlated with human adiponectin levels.[175] In the current study, the association of DPP4 activity with clinical group was not significant, which is a little unexpected given the strong negative correlation of prior laminitis status with adiponectin levels reported in Chapter 2. DPP4 activity was not associated with baseline insulin levels however it was positively associated with an increased insulin response to the oral sugar challenge. DPP4 activity was not associated with glucose and GLP-1 trajectories. The positive association of DPP4 activity with insulin and the negative association with adiponectin may be reflective of DPP4's role as an adipokine.[176] The negative association of DPP4 activity with age has also been previously reported in humans.[177,178] DPP4 activity was found to be higher in Welsh ponies than Morgan horses suggesting a possible genetic basis for

variation in DPP4 activity, although additive association testing of DPP4 activity with copy number of a single missense mutation in the *DPP4* gene was not significant.

3.5.2 Dynamic responses to the oral sugar test

This study utilized an an analytical approach that is not frequently used for studying variation in responses to an oral sugar challenge. Most diabetes researchers typically use a single summary statistic, area under the curve, when assessing variation in the oral glucose tolerance test. In this chapter, it was pointed out that response curves can be very different yet have the same area under the curve making it an adequate method for truly assessing variation in a response. This study found that an alternative approach, a longitudinal multilevel model, was able to test for differences in the trajectory of the response and identify factors that were significantly associated with variation in the trajectory that went undetected when assessing differences in the area under the curve. Mainstream use of longitudinal multilevel models for assessing oral glucose tolerance test variation potentially provides additional biologically relevant information and could improve conclusions drawn from studies that utilize the common oral glucose tolerance test.

Glucose trajectories

The initial rate of increase in glucose levels was found to decrease with age, a potential explanation for this finding may be age-related changes in intestinal glucose absorption.[179] Welsh ponies demonstrated an increased glucose response compared to Morgan horses potentially suggesting a genetic basis for variation in glucose response. Obesity status was associated with variation in the glucose trajectory, elevated baseline glucose values were observed in obese horses with no prior history of laminitis and increased glucose responses in obese horses with a prior history of laminitis replicating findings in human and rodent studies.[180,181] Dietary composition was associated with glucose trajectory variation with individuals on higher fiber diets having a reduced glycemic response. High fiber diets have been advocated in humans as a management tool for hyperglycemia.[182] Similar to previous equine studies, this study observed elevated glycemic responses in individuals consuming a higher water soluble carbohydrate diet.[183] Biochemical parameters, blood triglyceride and NEFA level, were positively

associated with glycemic responses. NEFA are known to both decrease insulin mediated glucose transport[184] and induce hepatic insulin resistance.[185] Insulin resistance results in failed suppression of NEFA during an oral sugar challenge resulting in increased gluconeogenesis.[186] Examining changes in NEFA levels in our study population has potential to shed additional light on the observed correlation of fasting NEFA levels with the glycemic response.

Insulin trajectories

Insulinemic responses were found to be significantly greater in Welsh ponies compared to Morgan horses which again suggest a genetic basis for variation in the equine response to an oral sugar challenge. Insulin trajectories also varied with clinical group status. Obese horses had higher baseline insulin values and insulinemic responses, although the difference only reached significance for obese horses with a prior history of laminitis. In the larger study population presented in Chapter 2, obesity and laminitis status were both found to be associated with post-OST insulin levels. Human studies have also demonstrated elevated fasting insulin and insulinemic responses to oral glucose with obesity and improvement of the insulinemic response with weight loss.[187,188] Similar to glucose trajectories, horses consuming a higher fiber diet had a reduced insulinemic response, providing support for dietary modulation as a dysinsulinemia management tool. Triglyceride and leptin were positively associated with insulinemic responses where as adiponectin was negatively associated similar to our findings in 75 minute post-OST insulin in the larger population presented in Chapter 2. Triglyceride, leptin, and adiponectin have all been shown to be associated with insulin resistance and insulin responses to oral glucose in humans,[189–191] however leptin is correlated with fat mass and the association with leptin has been shown not to occur independently of obesity.[192]

GLP-1 trajectories

Only a limited number of stallions (five) were included in the study population, but significant differences were observed in their active and total GLP-1 trajectories. Stallions exhibited reduced GLP-1 responses to oral sugar, however the baseline values were higher. Lower GLP-1 responses in the male gender has also been reported in humans.[90] GLP-1 trajectories varied with clinical group status. Total basal GLP-1 levels were

significantly lower in non-obese horses with a prior history of laminitis and also trended toward being lower in obese horses with and without a prior history of laminitis. Interestingly, obese horses with no prior history of laminitis had a significantly greater positive GLP-1 response where as obese horses with a history of laminitis had decreasing levels of GLP-1 throughout the oral sugar test. These findings indicate basal GLP-1 and GLP-1 secretory responses to oral glucose are reduced with features of EMS, similar to decreased GLP-1 secretion observed with features of human metabolic syndrome, obesity and insulin resistance.[168,193–195] Elevated NEFA levels were associated with decreased basal and secretory GLP-1. Human studies have demonstrated that pharmacologic suppression of NEFA levels results in a higher carbohydrate stimulated GLP-1 secretion.[193–195] Variation in basal GLP-1 and GLP-1 in response to oral sugar were correlated with diet nutrient composition. Higher starch and protein diets were associated with decreased GLP-1 secretory responses.

Variation in the proglucagon gene(*GCG*) which encodes GLP-1 was tested for association with GLP-1 trajectory. Cross-species alignment of 5' *GCG* flanking region nucleotides revealed an endogenous retrovirus long terminal repeat (ERV LTR) insertion located between what is referred to as the proglucagon gene upstream enhancer region (GUE) and the *GCG* promoter region. ERV LTRs naturally contain abundant transcriptional regulatory signals and often contribute to variation in mammalian gene expression, including tissue-specific expression, through donation of alternative promoters and enhancers. For example, an LTR sequence of a human endogenous retrovirus is involved in the tissue-specific expression of human salivary amylase.[196] The ERV LTR insertion upstream of three genes of the human amylase gene complex is correlated with a switch from pancreatic to parotid expression. Haplotype trend regression revealed an overall modest association with *GCG* haplotype. Interesting, haplotypes with higher basal GLP-1 levels tended to exhibit decreased GLP-1 responses to an oral sugar test. An interesting hypothetical explanation could involve GLP-1 expression in a tissue alternative to intestine due to alternative enhancers/promoters provided by the ERV LTR or disruption of the intestinal promoter/enhancers as a result of the insertion. Further experiments would be needed to test these hypotheses.

Correlation of trajectories

This study's use of a multivariate response longitudinal multilevel model allowed examination of the correlation of fasting values with initial rate of increase and deceleration both within and between each repeated response type. GLP-1 active and total values were strongly correlated for baseline values, initial rates of change, and deceleration rates were strongly correlated indicating minimal variation in GLP-1 active total ratios throughout the oral sugar test. Insulin and glucose were correlated for initial rates of change and deceleration however appreciable correlation was not evident between insulin and glucose with GLP-1 trajectories. If the equine incretrin response explained a large amount of the variation in the insulinemic response one would have expected to have observed stronger correlation of the GLP-1 and insulin trajectories. A possible explanation for the lack of observed correlation may be due to a differential relationship of GLP-1 secretion with the insulinemic response dependent on clinical group status which was not specifically tested in the unconditional longitudinal model used to determine correlation of the trajectories. In humans it has been shown that the incretin effect accounts for 70% of the variation in the insulin response to an oral sugar challenge in normal individuals but only 30% percent in individuals affected with type 2 diabetes.[197] Although, a recent study that more specifically measured the incretin effect with concomitant analysis of oral and intravenous glucose tests also reported a weak correlation of incretin hormone levels with the incretin effect (potentiation of oral glucose induced insulin secretion relative to intravenous glucose induced insulin secretion).[167]

Conclusions

Overall, this study replicated many of the findings associated with human incretin variation. Previously, it has been considered plausible that the pathophsyiology of equine hyperinsulinemia may be opposite of that which is believed to occur in humans, with horses manifesting chronic hyperinsulinemia as a result of an enhanced incretin response with insulin resistance occurring secondary to chronic hyperinsulinemia.[92,93] The results from our study do not support this hypothesis and suggest the pathophysiology of equine incretin response impairment may be similar to humans with insulin resistance playing a primary role. Additional studies that evaluate a horse's response to an oral

glucose stimulus compared with their response to an intravenous glucose stimulus are needed to further assess the incretin role in equine glucose intolerance.

Chapter 4

Improved linear mixed model of polygenic traits in populations with familial relatedness

4.1 Summary

Unexplained individual level variation and breed differences in metabolic phenotypes identified in Chapters 2 and 3 support the hypothesis that an underlying genetic susceptibility to EMS exists. Human MetS is a highly polygenic syndrome where numerous candidate genes have been identified. Whereas MetS associated variants are typically of small effect size; it was hypothesized that in EMS a small number of moderate to large effect loci contribute to variation in metabolic traits due to the fact that horse populations do not randomly mate and experience substantial selection pressure. 286 Morgan horses were genotyped on the Illumina SNP50 chip and imputed up to $>$800,000 SNPs to perform a genome wide association study (GWAS) to identify candidate genes for EMS. Additive genetic variance estimated from a genomic relationship matrix calculated from genotyped SNPs ("chip heritability") indicated that the 11 measured metabolic traits were moderately heritable. Yet initial genome-wide scans using standard linear mixed models failed to detect significant associations. Chapter 4 proposes an improved

linear mixed model for mapping polygenic traits in a population with familial relationships similar to that in many equine GWAS was developed and validated. The model incorporates a Bayesian variable selection method to rank SNPs and a stepwise feature selection process to determine the optimal SNPs to model the random polygenic effect, while including a random effect for each sampled herd or "familial cluster". The method was validated using the QTL-MAS 2010 dataset, and Morgan horse and Welsh pony height datasets, and demonstrated increased power while controlling the false positive rate.

4.2 Introduction

Association mapping of complex traits poses several challenges. First, complex traits are typically polygenic, meaning numerous genes contribute to the variability observed in a group of individuals measured for a particular trait. Statistical models testing the association of a trait with a single genetic marker will have reduced power to detect an association in comparison to a model that simultaneously considers additional loci contributing to variation of the trait. Power of the model will also vary dependent on the effect size of the causal variant being tested, the frequency of the genetic marker, and the degree of linkage disequilibrium (LD) between the causal variant and the genetic marker being tested.

Secondly, a homogeneous mapping population in terms of the level of relatedness between pairs of individuals is ideal, although unrealistic. Every population has varying levels of relatedness that may result in spurious associations due to allele frequency differences correlated with genetic background or spatial environment. The confounding structure and/or relatedness must be removed via sample design or accounted for in the statistical model.

The linear mixed model (LMM) effectively reduces the confounding effect of varying relatedness among individuals.[198–200] Fisher (1918) demonstrated phenotypic covariance between related individuals depends on their genetic relatedness for polygenic traits.[201] The more alleles individuals share, the more similar their phenotype. The classic Henderson mixed model (1984)[202] has been adapted to reduce the confounding effects of genetic background by estimating pairwise relatedness from all single

nucleotide polymorphisms (SNPs) to build a genetic relationship matrix (GRM) and model the random genotypic covariance effect on phenotype.[203] Relatedness serves as a proxy for allele sharing at causative loci. Even in ideal study populations without structure or familial relatedness, mixed linear models can increase power by implicitly conditioning on associated loci other than the candidate locus by reducing what would otherwise be considered by the model as random noise (i.e. unaccounted for loci of true effect would incorrectly be considered random noise resulting in larger error terms and reduced power).[204]

The standard LMM approach is based on single-locus tests combined with a diffuse, overall estimate of the genomic background based on all SNP markers. However, diffuse modeling of the polygenic term may not be appropriate for traits controlled by multiple loci of moderate to large effect. Recent LMM innovations aimed at improved modeling of the polygenic architecture in effort to increase power include modeling suspected causal loci as fixed effects[205] or building a GRM comprised of select SNPs associated with the phenotype.[206–209] An additional advantage of the select SNP GRM approach is increased computational efficiency due to using a lower-rank relatedness matrix (the number of SNPs used to determine the matrix is less than the number of individuals). Widmer et al[208] cautioned building a GRM based on selected SNPs that well predict the phenotype performed poorly at controlling the type I error rate in the presence of familial relatedness and recommended including a second GRM comprised of all SNPs to achieve adequate type I error control. Although, a drawback of this approach is increased computational time due to full-rank relatedness matrix inclusion.

In this chapter, a LMM approach featuring a novel select SNP GRM building algorithm with increased power and adequate type I error control will be presented and applied to the QTL mapping and marker assisted selection (QTL-MAS) 2010 workshop simulated dataset[210] and two real datasets featuring familial relatedness.

4.3 Methods

4.3.1 Description of simulated data

The 2010 QTL-MAS[210] simulated dataset features familial relationships (half-sibs, full-sibs) and is comprised of 3226 individuals from five consecutive generations (F0-F4). There were 20 founders: 5 males and 15 females. The pedigree structure was created assuming each female mates once and gives birth to approximately 30 progeny. Almost non-overlapping generations were created (the parents from every next generation were selected at random mostly from the current generation). Each of the 2326 individuals in generation F0 to F3 has phenotypic records on a quantitative trait and were included in the analysis. A genome consisting of 10,031 biallelic SNPs on 5 chromosomes were simulated without any missing data and genotyping error. 37 SNPs were selected as causal variants and the genotypes for 28 QTL were removed from the dataset. The quantitative trait was simulated with heritability of 0.39. SNPs with a minor allele frequency less than 0.05 were excluded from the analysis.

4.3.2 Description of real data

The Morgan horse height data consists of 286 individuals sampled from 54 farms located throughout the United States and a single Canadian farm. The horses were genotyped on on the Illumina SNP50 array and imputed up to ~2 million genotypes from a cosmopolitan reference population comprised of 384 individuals representing 30+ breeds, including 43 horses from the Morgan breed using Beagle 4.0 software.[211,212] Genotypes were filtered for MAF < 0.05, and Hardy-Weinberg disequilibrium p-values < 0.00001 yielding 877,608 SNPs. The heritability estimate determined using the GEMMA software[213] adjusted for age and sex covariates and estimated from a centered GRM comprised of all 845,234 autosomal SNPs was 1.00 (standard error: 0.00).

The Welsh pony height data consists of 234 individuals sampled from 27 US farms genotyped on the Affymetrix 670K array. Genotypes were filtered for SNP missingness > 0.05 (remaining missing genotypes were imputed using Beagle 4.0 software[211,212]), MAF < 0.05, and Hardy-Weinberg disequilibrium p-values < 0.00001 yielding 394,880 SNPs. The heritability estimate determined using the GEMMA software[213] adjusted for age and sex covariates and estimated from a centered GRM comprised of all 379,604

autosomal SNPs was 0.95 (standard error: 0.03).

4.3.3 The linear mixed model

The linear mixed model decomposes the phenotype ($\mathbf{y}$) into the phenotype mean ($\boldsymbol{\mu}$), fixed covariate ($\boldsymbol{\beta}_{\text{cov}}$) and fixed genetic marker ($\boldsymbol{\beta}_{\text{genetic marker}}$) effects, random genetic effect ($\mathbf{u}$), and residual ($\boldsymbol{\epsilon}$) as follows:

$$\begin{aligned} \boldsymbol{y} = \boldsymbol{1_n\mu} + \boldsymbol{W\beta}_{\text{cov}} + \boldsymbol{x\beta}_{\text{genetic marker}} + \boldsymbol{u} + \boldsymbol{\epsilon}; \\ \boldsymbol{u} \sim \boldsymbol{MVN_n(0, \lambda\tau^{-1}K)}, \\ \boldsymbol{\epsilon} \sim \boldsymbol{MVN_n(0, \tau^{-1}I_n)}, \end{aligned} \tag{4.1}$$

where $\mathbf{1n}$ is a n-vector of ones; $\boldsymbol{\mu}$ is a scalar representing the intercept; $\mathbf{W}$=(w$_{\mathbf{1}}$,...,w$_{\boldsymbol{c}}$) is a n x c matrix of covariates; $\boldsymbol{\beta}_{\text{cov}}$ is a c-vector of covariate effect sizes; $\mathbf{x}$ is a n-vector of marker genotypes; $\boldsymbol{\beta}_{\text{genetic marker}}$ is the effect size of the marker; $\mathbf{u}$ is a n-vector of unknown random polygenic effects; $\boldsymbol{\epsilon}$ is a n-vector of random residual effects; $\boldsymbol{\tau^{-1}}$ is the variance of the residual errors; $\boldsymbol{\lambda}$ is the ratio between the two variance components; $\mathbf{K}$ is a known n x n relatedness matrix with element $\mathbf{K}_{\boldsymbol{ij}}$ (i,j=1,2,...,n) calculated from genetic markers and $\mathbf{I}_n$ is a n x n identity matrix. MVN$_n$ denotes the n-dimensional multivariate normal distribution. Solving equation (4.3) involves determining all the unknown parameters under which the observations ($\mathbf{y}$) have the maximum likelihood, defined as follows:

$$\boldsymbol{L(y|\mu, \beta}_{\text{cov}}\boldsymbol{, \beta}_{\text{genetic marker}}\boldsymbol{, \lambda, \tau^{-1})}. \tag{4.2}$$

The alternative hypothesis $H_{\mathbf{1}} : \boldsymbol{\beta}_{\text{genetic marker}} \neq \mathbf{0}$ is tested for each genetic marker (SNP) in turn against the null hypothesis $H_{\mathbf{0}} : \boldsymbol{\beta}_{\text{genetic marker}} = \mathbf{0}$.

4.3.4 The bayesian sparse linear mixed model

The bayesian sparse linear mixed model (BSLMM), as described in Zhou et al 2013,[207] is a hybrid between a linear mixed model (assumes all genetic markers have an effect) and bayesian variable selection regression (BVSR) that assumes genetic marker effects come from a mixture of two normal distributions where all genetic markers have a small effect while a portion of the genetic markers have additional larger effects and has the

following form:

$$\begin{aligned} \boldsymbol{y} &= \mathbf{1}_n\boldsymbol{\mu} + \boldsymbol{X\beta} + \boldsymbol{u} + \boldsymbol{\epsilon}; \\ \boldsymbol{\beta_i} &\sim \boldsymbol{\pi N(0, \sigma_a^2 \tau^{-1})} + \boldsymbol{(1-\pi)\delta_0}, \\ \boldsymbol{u} &\sim \boldsymbol{MVN_n(0, \sigma_b^2 \tau^{-1} K)}, \\ \boldsymbol{\epsilon} &\sim \boldsymbol{MVN_n(0, \tau^{-1} I_n)}, \end{aligned} \tag{4.3}$$

$\mathbf{X}$ is a n x p matrix of genotypes measured on n individuals at p genetic markers, $\boldsymbol{\beta}$ is the corresponding p-vector of the genetic marker effects, $\boldsymbol{\pi}$ is the proportion of genetic markers with large effects, $\boldsymbol{\delta_0}$ denotes a point mass at zero (no large effect), $\boldsymbol{\sigma_a}$ is the variance of the large SNP effects, $\boldsymbol{\sigma_b}$ is the variance of the small SNP effects, and other parameters are the same as defined in the standard linear mixed model in the previous section. Setting $\boldsymbol{\pi} = \mathbf{0}$ is equivalent to running a LMM whereas setting $\boldsymbol{\sigma_b} = \mathbf{0}$ is equivalent to BVSR. When $\mathbf{K} = \mathbf{X}\mathbf{X}^T/p$ (centered relationship matrix), the SNP effect sizes are decomposed into two parts: 1) $\boldsymbol{\alpha}$ captures the small effect that all SNPS have and 2) $\boldsymbol{\beta}$ capture the additional large effects of some SNPs. $\mathbf{u} = \mathbf{X}\boldsymbol{\alpha}$ is the combined effect of all small effects. The total effect for a given SNP is $\boldsymbol{\alpha_i} + \boldsymbol{\beta_i}$. Hyper-parameters $\boldsymbol{\mu, \tau, \pi, \sigma_a, \sigma_b}$ are estimated in a Bayesian framework by specifying prior distributions for the parameters and using Markov chain Monte Carlo (MCMC) to obtain approximate samples from their posterior distribution given the observed data.

4.3.5 Description of improved linear mixed model algorithm for mapping polygenic traits in populations with familial relatedness

The improved LMM algorithm is a basic 3 step process:

- **Step 1:** Bin each chromosome into 500 kilobase (KB) segments and select the most influential SNP in addition to the two immediately adjacent SNPs to represent each bin (inclusion of adjacent SNPs demonstrated superior control over the false positive rate in comparison to including only a single SNP to represent a bin in the simulated dataset (see Table 4.1)).

- **Step 2:** Perform stepwise feature (bin) selection to build the select SNP GRM.

- **Step 3:** Perform LMM GWAS with select SNP GRM in place of an all SNP GRM.

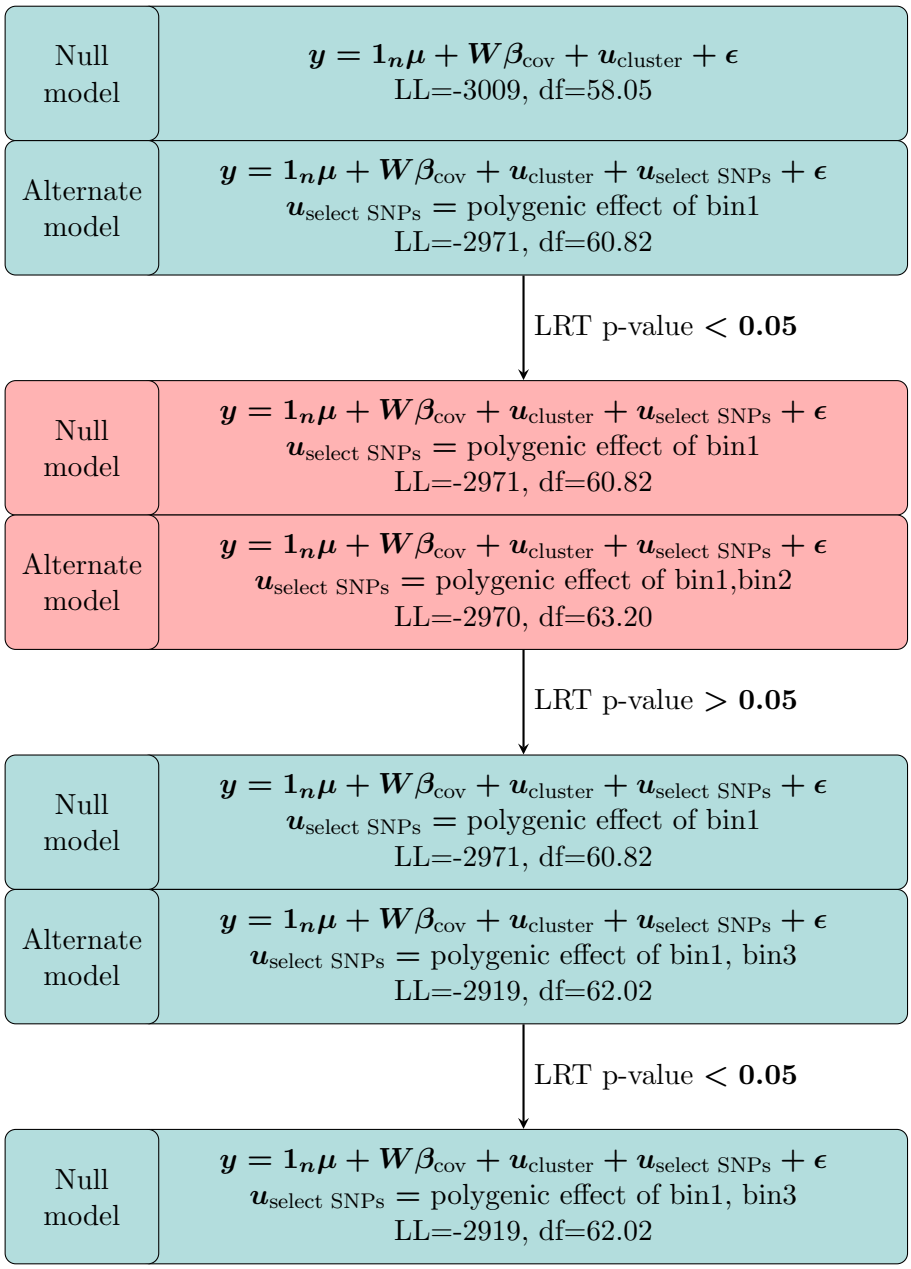
In Step 1, BSLMM was performed to rank SNPs based on model frequency (the proportion of iterations that the SNP was included in the subset of markers deemed to have a large effect). The SNP with the highest BSLMM model frequency in each 500KB bin was deemed the most influential and selected to represent the bin along with the two immediately adjacent SNPs. The GEMMA software[207] was used to perform the BSLMM. The phenotypes were first adjusted for covariates (sex in the simulated dataset; age and sex in the real dataset) prior to the BSLMM analysis as the software does not incorporate inclusion of covariates. Covariate effects were estimated from a linear mixed effect model including a random intercept for the clusters. Adjusting for covariates prior to the BSLMM assumes the covariate effects are unrelated to genetic effects which may not be a correct assumption in all situations. 550,000 MCMC iterations were performed (including 50,000 burn-in iterations). The number of iterations selected was influenced by computation speed. Variation in results dependent on number of iterations was not assessed in the current study and is left to the user to determine the appropriate balance between computational speed and accuracy.

Step 2 involves a stepwise feature selection model building process to determine the ideal bins to be included in the select SNP GRM. A schematic illustrating Step 2 is presented in Figure 4.1. First, a null model which includes fixed effect covariates and a random intercept term for each genetic cluster (full-sib families in the simulated data; farm in the real datasets) is compared with an alternative model that includes an additional polygenic random effect determined from a select SNP GRM calculated from the SNPs representing the top ranked bin from the BSLMM. A second GRM composed of all SNPs, similar to Widmer et al's[208] recommendation was considered in place of the random intercept, however there was no gain in power or false positive control (see Table 4.1). Therefore, due to improved computational speed, the background random intercept was elected over a second GRM composed of all SNPs. A likelihood ratio test is performed to determine if inclusion of SNPs representing the bin into the select SNP GRM significantly improve the model. The degrees of freedom are determined from the difference in the effective degrees of freedom of each model.[214,215] If inclusion of the bin is significant, the alternative model becomes the new null model and is tested against

an alternative model that is the same as the new null model with the exception that the SNPs representing the next highest ranked bin are also included in the calculation of the select SNP GRM. Under the condition where the alternative model is not significantly improved over the null model, the SNPs representing the current bin are removed from the select SNP GRM in exchange for the SNPs representing the next ranked bin. The process continues until all bins have been tested for inclusion in the select SNP GRM.

After completion of Step 2, the select SNP GRM is built and ready for inclusion in Step 3, the improved LMM GWAS. During the testing of a SNP in the final improved LMM GWAS, SNPs within 1MB of a test SNP are removed from the select SNP GRM to avoid proximal contamination.[206,216] Exclusion of SNPs from the select SNP GRM based on LD with test SNPs, similar to Wang et al.[209] was considered however false positive control was reduced and computational time increased (data not shown). The final GWAS model includes covariates and a random intercept for each genetic cluster in addition to the select SNP GRM. The FaST-LMM software is used to estimate parameters in steps 2 and 3 by maximizing the log-likelihood using the algorithms of Lippert et al.[206,217] and Widmer et al.[208]

Figure 4.1: Schematic illustration of feature selection process to build select SNP genomic relationship matrix for inclusion in improved linear mixed model GWAS.



4.3.6 Validation of the improved linear mixed model

For the QTL-MAS simulated data, the comparison criteria was similar to that used by the 2010 QTL-MAS workshop; a true QTL was considered mapped if one or more of the deemed QTL positions were within 1 MB distance from the simulated QTL. If two or more of the deemed QTL positions were within 1 Mb distance from a simulated QTL, they were considered to map the same true QTL. Reported positions exceeding 1MB from the true QTL were considered to be false positives. For both the standard and improved LMM, Wald test statistics with p-values $<\mathbf{1x10^{-5}}$ were considered statistically significant. The QTL positions were determined by starting with the position of the top ranked statistically significant SNPs and removing any other statistically significant SNPs in linkage disequilibrium ($\boldsymbol{r^2 > 0.2}$) with the top ranked SNP, the process was repeated for the 2nd top ranked and continued for all remaining statistically significant SNPs.

For the horse height datasets, the true QTL locations are unknown in the real datasets therefore it is not possible to report the number of detected true and false positives, however replication in independent datasets is considered the gold standard for validation of GWAS hits and therefore was used as a method of validating the improved algorithm in the real datasets. Both replication of previously reported horse loci and height loci in other species were considered. A randomized permutation test was used to formerly test for enrichment of human height QTL by evaluating how likely it was to produce the number of QTL observed within 2MB of an equine ortholog to the human gene nearest one of the 180 statistically significant SNPs identified in a human height GWAS[218] by chance. SNPs in a quantity equal to the observed number of horse height QTL detected were selected at random and the number of SNPs located within 2MB of an equine ortholog to a human height gene were determined. 10,000 random tests were performed.

4.4 Results

4.4.1 Application and optimization of the improved linear mixed model using simulated data

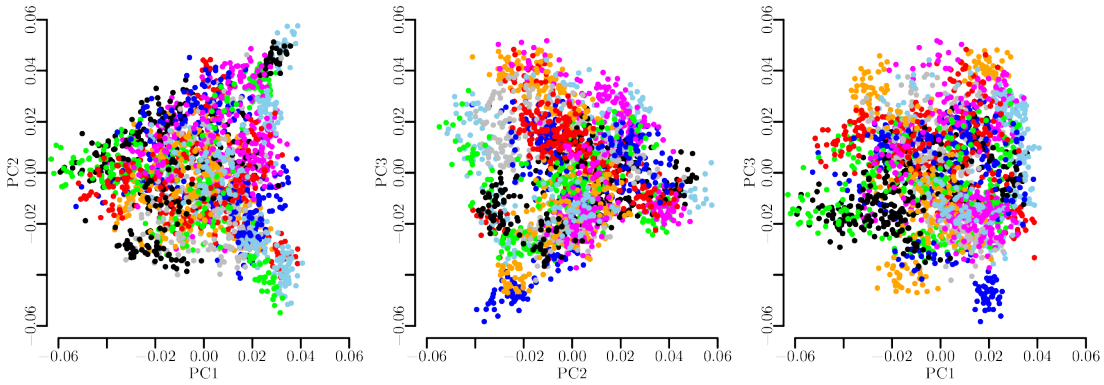
The improved linear mixed model (LMM) described above was applied to the simulated 2010 QTL-MAS dataset.[210] Principal component analysis (PCA) was performed and the familial structure of the data is illustrated in Figure 4.2a. Individuals from the same familial cluster share the same color and it is easy to see from the figure that colors are not randomly distributed. As expected, individuals from the same familial cluster are more genetically similar. In an effort to quantify the differences in average allele sharing for pairs of individuals sampled from the same familial cluster versus pairs of individuals from different familial clusters, the mean identity by state (IBS) was calculated for pairs of individuals from the same familial cluster versus pairs of individuals from different clusters. Mean IBS for pairs sampled from the same familial cluster was 0.82 in comparison to 0.69 in individuals sampled from different familial clusters. Figure 4.3a illustrates the differences in the distribution of IBS values for pairs of individuals belonging to the same familial cluster versus different familial clusters.

Wald tests from the improved LMM were compared with the standard LMM implemented in GEMMA[213] in terms of power (number of mapped true QTL) and number of falsely reported QTLs. Manhattan plots for the standard LMM and improved LMM for the simulated data are presented in Figure 4.4. At the suggestive significance level (p-values $< \mathbf{1x10^{-5}}$), the standard LMM mapped 6 true QTL with zero false discoveries (genome-wide Bonferonni corrected p-values yielded 5 true QTL detections and zero false discoveries). The improved LMM detected 13 true QTL along with a single false discovery at the suggestive significance threshold (12 true positives and a single false positive detected at the genome-wide Bonferonni threshold).

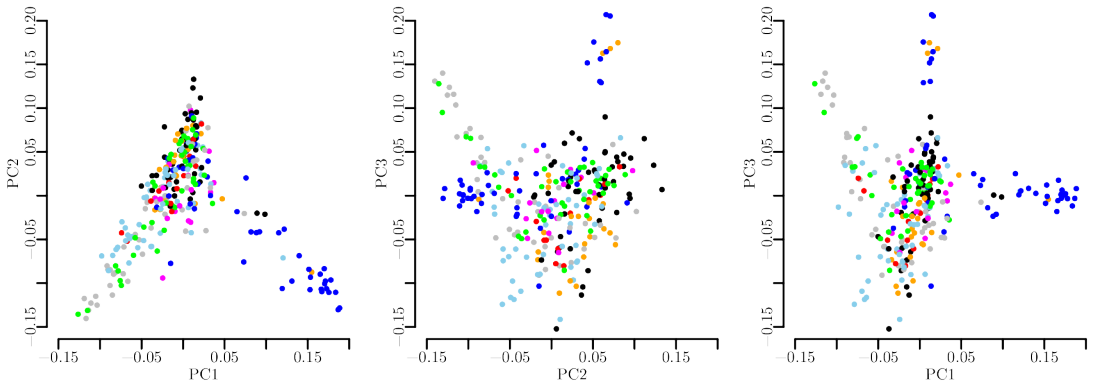
Utilizing a Bayesian model to initially rank the SNPs in the improved LMM algorithm performed superior to using an initial standard LMM to rank the SNPs prior to performing feature selection for inclusion in the select SNP GRM in the final GWAS (see Table 4.1). Manhattan plots of the initial BSLMM posterior inclusion probabilities (model frequency) used to initially rank the SNPs are presented in Appendix figure C.3 and a manhattan plot of GWAS results from an improved LMM model that substitutes

Figure 4.2: Familial relatedness. Principal components 1-3 for a) QTL-MAS simulated dataset, b) Morgan horse real height dataset, and c) Welsh pony real height dataset. Individuals sampled from the same family or farm are identified by a similar color (Note: colors are repeated for different clusters, an enlarged version of the horse PCA plots is presented in Figures C.1 and C.2 with numeric representation of different farms).

(a)



(b)



(c)

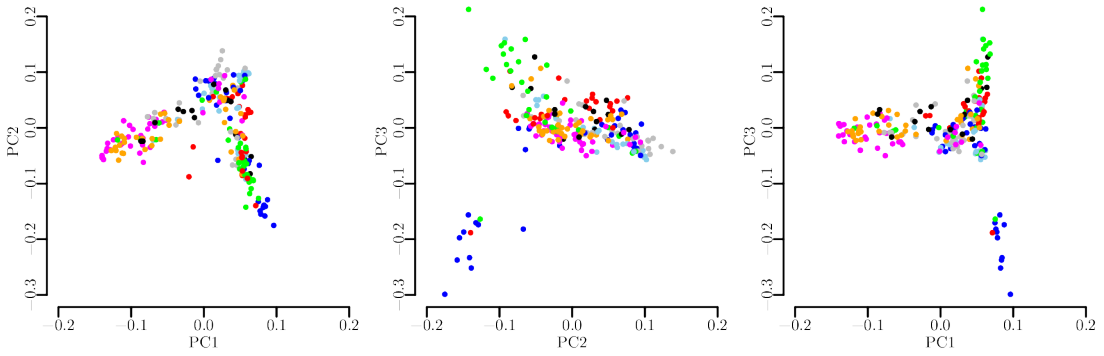
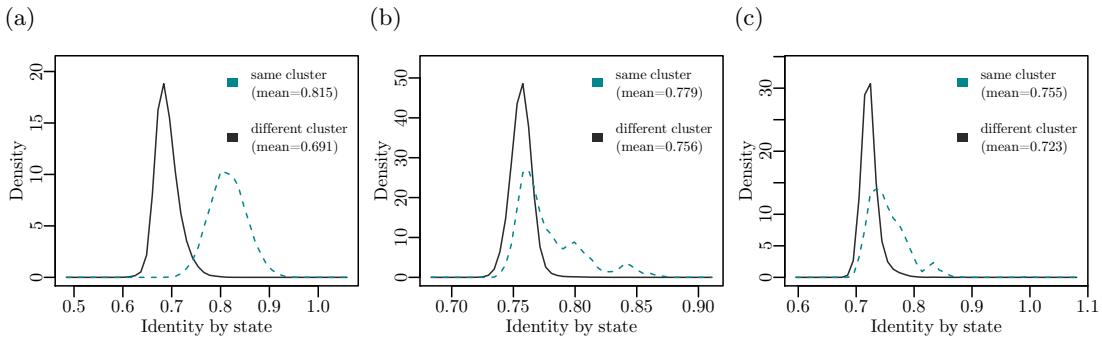


Figure 4.3: Illustration of IBS distribution difference for individuals sampled from the same cluster (farm) versus different clusters (farms) for the a) QTL-MAS simulated dataset, b) Morgan horse real height dataset, and c) Welsh pony real height dataset.



a LMM for the BSLMM in the first step to rank the SNPs is presented in Figure 4.5. A smaller bin size, 200KB instead of 500KB, was also evaluated and yielded similar results, see Appendix Figure C.4.

Table 4.1: Optimization of improved LMM.

Initial	top SNP representing bin		top SNP + adjacent SNPs	
SNP ranking model	True QTL detected	False positives	True QTL detected	False positives
LMM	10	5	9	2
BSLMM	13	4	13	1

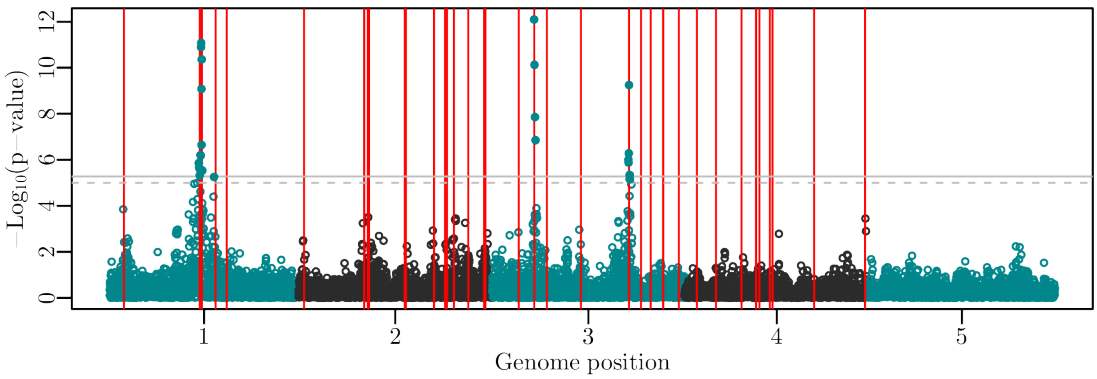
Note: bin size=500KB; LMM=standard linear mixed model, BSLMM=Bayesian sparse linear mixed model.

4.4.2 Application of the improved linear mixed model to equine datasets

The performance of the improved LMM algorithm was also compared to the standard LMM algorithm in 2 real datasets. Both datasets feature height, a classic polygenic trait) sampled in two different equine populations, the Morgan horse and the Welsh pony. These datasets are also characterized by familial relationships. Horses and ponies sampled from the same farms were genetically more similar to each other than individuals sampled from different farms. Principal component analysis in Figures 4.2b,c

Figure 4.4: Manhattan plot displaying the a) standard and b) improved linear mixed model results for the simulated QTL-MAS dataset. Red vertical lines indicate the location of true QTL. Genome-wide p-value $<$ **0.05** (solid grey horizontal line) and suggestive (dashed grey horizontal line) thresholds are also shown.

(a)



(b)

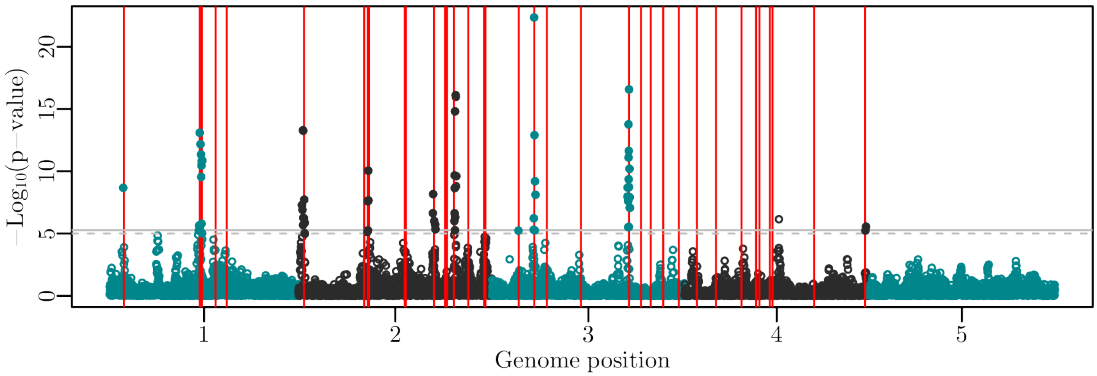
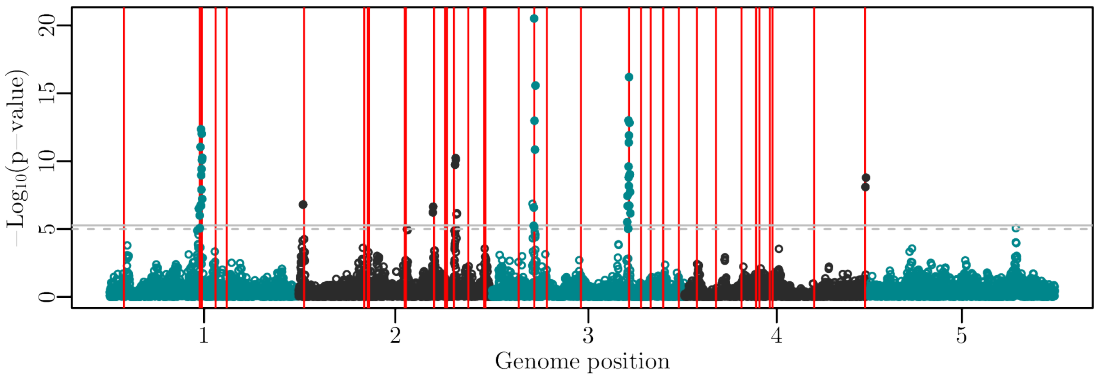


Figure 4.5: Manhattan plot displaying the improved linear mixed model results for the simulated QTL-MAS dataset using an initial standard linear mixed model to rank SNP. Red vertical lines indicate the location of true QTL. Genome-wide p-value **<0.05** (solid grey horizontal line) and suggestive (dashed grey horizontal line) thresholds are also shown.



illustrate the family structure present in these datasets. Pairs of individuals from the same farm share a similar color and similar to the simulated data it is easy to observe the non-random distribution of colors. Again, similar to the simulated data, mean IBS is higher in pairs of individuals sampled from the same farm in comparison to pairs of individuals sampled from different farms. Illustration of differences in the average proportion of alleles identical by state for individuals sampled from the same farm versus different genetic clusters is presented in Figure 4.3b,c.

Manhattan plots of GWAS results from the standard and improved LMMs are presented in Figures 4.6 and 4.7 for the Morgan horses and Welsh ponies, respectively. Since the true QTL locations are unknown in the real datasets it is not possible to report the number of detected true and false positives, however replication in independent datasets is considered the gold standard for validation of GWAS hits and therefore was used as a method of validating the improved algorithm in the real datasets. The top ranked QTL in both the Morgan horse and Welsh pony data for both the standard and improved LMM models are located within 1MB of genes (*ZFAT* and *HMGA2*, respectively) previously identified as QTL in horse,[219–221] human,[218] and cattle[222] GWAS and candidate gene studies. P-values for the *ZFAT* locus in the standard and improved LMM in the Morgan horse population were $\mathbf{2.21x10^{-7}}$ and $\mathbf{1.44x10^{-18}}$, respectively,

while p-values for the *HMGA2* locus in the standard and improved LMM in the Welsh population were $\mathbf{1.80x10^{-26}}$ and $\mathbf{7.74x10^{-87}}$, respectively.

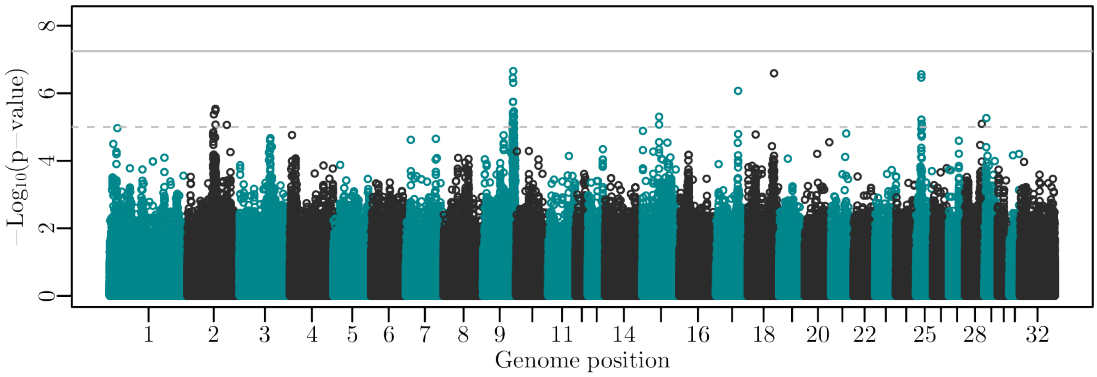
Similar to evaluations in the simulated data, the improved LMM identified substantially more QTL than the standard LMM for the real datasets. In the Morgan horses, 8 genome-wide significant loci (including 4 human candidate loci) and 13 suggestive loci (including 4 human candidate loci) were identified by the improved LMM (Table 4.2) in comparison to zero genome-wide significant loci and 10 suggestive loci (including 3 human candidate loci) identified by the standard LMM. In the Welsh ponies, 9 genome-wide significant loci (including 4 human candidate loci) and 6 suggestive loci (including 3 human candidate loci) were identified by the improved LMM (Table 4.3) in comparison to 2 genome-wide significant loci (including 1 human candidate locus) and 2 suggestive loci (including 1 human candidate locus) identified by the standard LMM. The top ranked loci overlapped for the improved and standard LMMs in both the Morgan horse and Welsh ponies. In addition, the *ZNF462* locus was detected at the genome-wide and suggestive significance level in the Morgan horse and Welsh pony, respectively, providing further support of *ZNF462* being a true QTL for height variation in horses. The *IPPK* and *GPR126* loci detected in the Morgan horse population have also been reported as QTL for cattle height in addition to human height.[222] Additional candidate genes for horse height may localize to the identified QTL that did not harbor a human candidate height gene and these loci are provided in Tables 4.2 to 4.3 for the interested reader. A search for potential height candidate genes was not conducted since mapping height QTL was not the objective of the current study.

A randomized permutation test was used to formerly test for enrichment of human height QTL by evaluating how likely it was to produce the number of QTL observed within 2MB of an equine ortholog to the human gene nearest one of the 180 statistically significant SNPs identified in a human height GWAS[218] by chance. Only 20% and 9% of the randomized permutation tests in Morgan horses and Welsh pony data, respectively, produced the number of QTL equal to or exceeding the results observed at the suggestive (p-values $< \mathbf{1x10^{-5}}$) significance level and only 14% and 22%, at the Bonferonni genome-wide significance level. Therefore, it is unlikely that the results reported here for height in Morgan horse and Welsh ponies arose by chance. In addition, evaluation of Q-Q plots in Figure 4.8a-c illustrate the observed test statistic distribution with the

expected null distribution and demonstrate effective control of the type 1 error rate with genomic inflation factors less than one for the simulated and real datasets.

Figure 4.6: Manhattan plot displaying the a) standard and b) improved linear mixed model results for the Morgan horse real height dataset. Genome-wide p-value$<\mathbf{5.70x10^{-8}}$ (solid grey horizontal line) and suggestive p-value$<\mathbf{1x10^{-5}}$(dashed grey horizontal line) thresholds are also shown.

(a)



(b)

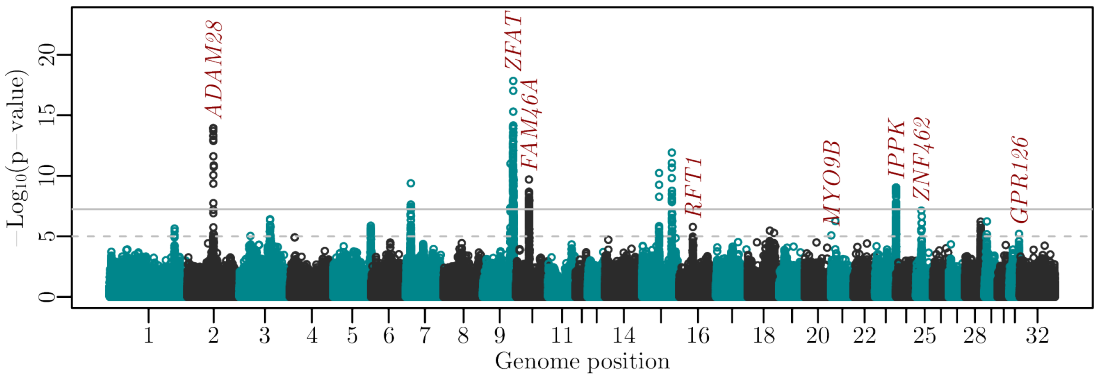


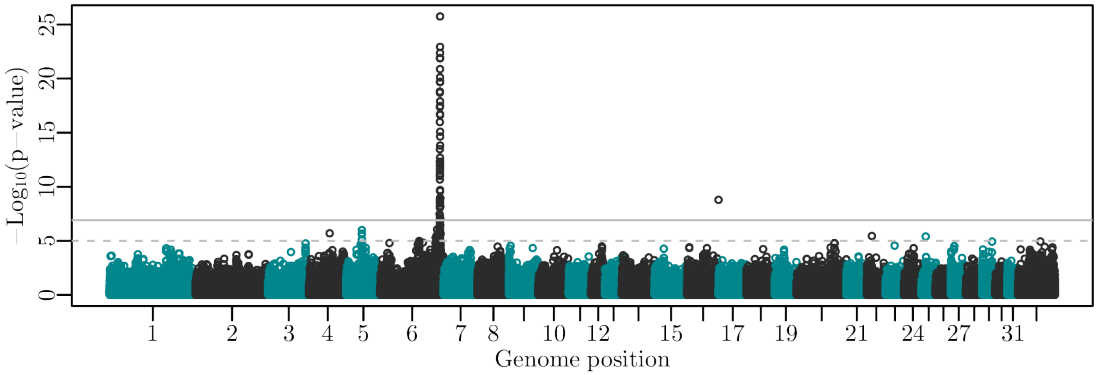
Table 4.2: Height QTL identified in Morgan horse population by the improved linear mixed model. Bonferonni threshold: p-value$<\mathbf{5.70x10^{-8}}$ Suggestive threshold: p-value$<\mathbf{1x10^{-5}}$

chromosome	position	beta estimate	standard error	p-value	human height candidate gene
9	75674706	0.259	0.029	1.44E-18	*ZFAT*
2	54258131	-0.237	0.031	1.17E-14	*ADAM28*
15	73417660	-0.194	0.027	1.21E-12	
9	69065870	-0.186	0.027	9.91E-12	
15	41710938	-0.193	0.029	5.88E-11	
10	34921358	-0.172	0.027	2.01E-10	*FAM46A*
7	11470079	-0.175	0.028	4.14E-10	
23	54928132	-0.162	0.027	9.35E-10	*IPPK*
25	12606038	-0.157	0.029	7.31E-08	*ZNF462*
3	73225735	-0.138	0.027	3.92E-07	
21	12777073	-0.135	0.027	5.48E-07	
29	6735168	0.126	0.025	5.84E-07	
28	36628887	-0.133	0.027	6.11E-07	
5	98798764	0.133	0.027	1.33E-06	
16	33334177	0.142	0.03	1.69E-06	*RFT1*
1	152515464	0.131	0.028	2.27E-06	
18	58850432	-0.12	0.026	3.40E-06	
18	69763301	-0.127	0.028	5.21E-06	
31	22766464	-0.13	0.029	6.30E-06	*GPR126*
21	586856	0.112	0.025	8.27E-06	*MYO9B*
3	23442441	-0.125	0.028	9.29E-06	

Note: *ZFAT* is a previously identified QTL for height in horses.[220,221]

Figure 4.7: Manhattan plot displaying the a) standard and b) improved linear mixed model results for the Welsh pony real height dataset. Genome-wide p-value$<\mathbf{1.27x10^{-7}}$ (solid grey horizontal line) and suggestive p-value$<\mathbf{1x10^{-5}}$(dashed grey horizontal line) thresholds are also shown.

(a)



(b)

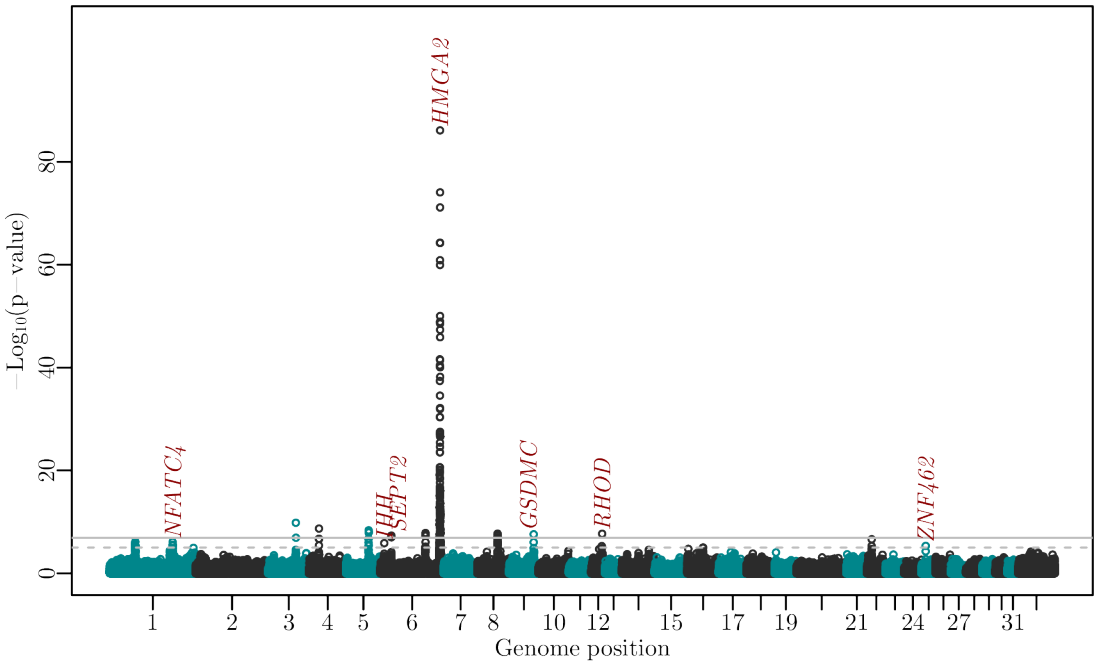


Table 4.3: Height QTL identified in Welsh pony population by the improved linear mixed model. Bonferonni threshold: p-value$<\mathbf{1.27x10^{-7}}$ Suggestive threshold: p-value$<\mathbf{1x10^{-5}}$

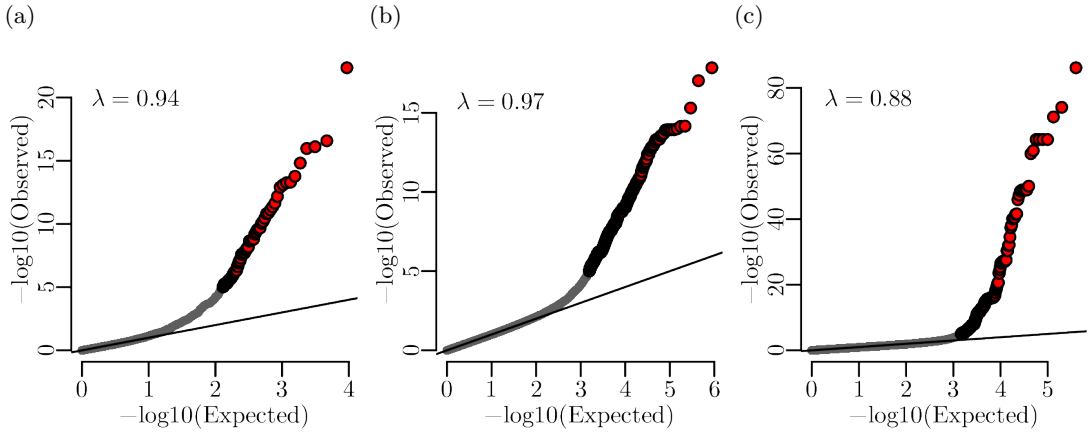
chromosome	position	beta estimate	standard error	p-value	human height candidate gene
6	81413254	0.466	0.024	7.74E-87	*HMGA2*
3	81805228	0.122	0.019	1.50E-10	
4	26707485	-0.116	0.019	2.04E-09	
5	68039446	0.108	0.018	4.47E-09	
6	66913666	0.112	0.02	1.43E-08	
8	60233394	-0.119	0.021	1.94E-08	
12	26848333	-0.126	0.022	2.08E-08	*RHOD*
9	70148789	0.12	0.022	2.64E-08	*GSDMC*
6	28058021	0.107	0.02	4.74E-08	*SEPT2*
22	11893884	0.097	0.019	2.43E-07	
1	75767477	0.09	0.018	7.31E-07	
1	161563724	-0.105	0.021	8.47E-07	*NFATC4*
6	9239939	-0.092	0.019	1.41E-06	*IHH*
25	12661868	0.105	0.023	4.67E-06	*ZNF462*
16	46179071	0.091	0.021	9.63E-06	

Note: *HMGA2* is a previously identified QTL for height in ponies.[219,220]

4.4.3 Variance decomposition during the stepwise feature SNP selection process

During step 2 of the improved LMM (stepwise feature SNP selection process) a majority of the bin representative SNPs selected for inclusion in the select SNP GRM are chosen in the beginning of the feature selection process with very with SNPs representing lower ranked bins being included in the select SNP GRM, see Appendix Figure C.5. This is not surprising given SNPs were initially ranked by the BSLMM based on their model inclusion probability of having more than a small effect on the trait. Also, the cluster level variance is reduced to zero as SNPs are selected that presumably explain phenotypic variation between the clusters, i.e. correlation of the cluster allele frequency with the phenotype. A total of 38 bins were selected for the simulated dataset and explained 47% of the total phenotypic variance. In the Morgan horse data, 41 bins were selected,

Figure 4.8: Q-Q plot of observed p-values against the expected p-values for a) QTL-MAS simulated dataset, b) Morgan horse real dataset, and c) Welsh pony real dataset.



explaining 85% of height variation and in the Welsh pony data 26 bins were selected explaining 91% of height variation.

4.5 Discussion

Linear mixed models (LMM) are regarded as the method of choice for association testing in genome-wide association studies (GWAS) as they account for population structure, familial and cryptic relatedness while achieving increased statistical power by jointly modeling all of the genotyped markers.[198,199,203,205,206,213,216,223–226] However, the standard LMM assumes that all genetic markers are causal (the so-called "infinitesimal model') with small effect sizes drawn from an independent Gaussian distribution which is unlikely to be true, especially in domesticated animal populations. Recently, efforts to more accurately model a non-infinitesimal genetic architecture have been proposed. Segura et al.[205] proposed identifying and conditioning out loci of large effect as fixed effects while others have recommended applying the mixed model to only a selected subset of phenotype associated markers.[208,209,214,216,224,227] Widmer et al.[208] cautioned using only phenotype associated markers in the LMM genomic relationship matrix (GRM) led to insufficient control of the type I error rate in populations with

familial relatedness, therefore recommending inclusion of a second GRM composed of all genetic markers to correct for familial structure while still maintaining statistical power to detect causal variants with the caveat that inclusion of a second GRM composed of all SNPs is much more computationally intensive. Chapter 4 described a new algorithm that builds upon these recent advances while maintaining computational feasibility and demonstrated the increased power of the improved LMM algorithm over the standard LMM while still maintaining control of the false positive rate.

The improved LMM algorithm is an advance for multiple reasons. First, the improved LMM more accurately models the true genetic architecture by applying a Bayesian method to initially rank SNPs combined with stepwise feature selection to select an optimal subset of SNPs for inclusion in the GRM. Recently, Bayesian methods have been developed that apply non-infinitesimal model assumptions to produce improved genetic prediction accuracy.[207,228] The improved LMM extends this concept to association testing but still provides the traditional GWAS frequentist test statistic.

Secondly, the improved LMM harnesses the power of phenotype specific variant inclusion in the GRM while avoiding the computationally intensive inclusion of a second GRM composed of all SNPs to correct for familial relationships. The improved LMM simply includes a random intercept effect for each genetic cluster of individuals while selecting feature SNPs for inclusion in the select SNP GRM and when calculating the GWAS test statistics. Individuals clustered in similar environments are often more similar genetically, especially in domesticated animal populations. The random intercept effectively captures shared environment and genetic background effects when attempting to select SNPs tagging causal variants that explain phenotypic differences in individuals that share a common environment or genetic background, simultaneously increasing power and correcting for non-independent observations. In situations where family structure is not correlated with shared environment it may be necessary to include a second GRM determined from all autosomal SNPs to adequately model the family structure and prevent inflation of test statistics.

Finally, the improved LMM algorithm presented in this chapter shares many similarities with the BOLT-LMM method proposed by Loh et al.[229] in terms of employing a Bayesian framework yet providing a frequentist test statistic, however the BOLT-LMM method recommends randomly ascertained observations and has not been studied in

datasets with family structure. Compared with currently available advances in LMM for GWAS, the improved LMM described in this work offers increased statistical power while maintaining control of the false positive rate in populations with family structure in a computationally efficient manner.

Chapter 5

Genome wide association study of equine metabolic trait variation

5.1 Summary

Chapter 5 utilizes the improved linear mixed model described in Chapter 4 to identify genetic variants influencing metabolic syndrome in horses. Using this improved linear mixed model, 76 suggestive and 17 genome-wide significant candidate loci were identified for the 11 metabolic traits in the 286 Morgan horse cohort. Candidate genes substantially overlapped with human MetS candidate genes including: VEGFA, NRXN3, GRIK2, and TRIB2. Other interesting candidate genes included ISL, which encodes insulin enhancer protein that is thought to play an important role in regulating insulin gene expression; and AHR which encodes the aryl hydrocarbon receptor, a ligand activated transcription factor known to bind endocrine disrupting chemicals such as polycyclic aromatic hydrocarbons and dioxins. AHR is an interesting candidate gene given the potential role of endocrine disrupting chemical in the pathophysiology of MetS, and unexplained sources of farm level variation in Chapter 2. The shared candidate genes for metabolic syndrome in humans and horses suggests similar underlying pathophysiological mechanisms and provides opportunity for exploring similar preventative and therapeutic management strategies.

5.2 Introduction

Equine metabolic syndrome (EMS) is a complex disease influenced by genetic and environmental factors and their interactions. This study presents findings from a GWAS designed to identify genetic variants influencing metabolic syndrome in horses. Eleven metabolic phenotypes described in Chapter 2 representing not only obesity, but also biological processes associated with obesity were tested for association with a dense panel of 877,608 single nucleotide polymorphisms (SNP)s in a population of 286 Morgan horses sampled from 55 herds. Horses sampled from the same farms exhibited similar phenotypes and increased genetic similarity, thus the improved linear mixed model (LMM) described in Chapter 4 was utilized to test for associations while controlling for shared environment and polygenic background. Loci identified to be associated with equine metabolic phenotypes (p-value$< \mathbf{1x10^{-5}}$) were investigated for overlap with loci identified in human GWAS of obesity and its comorbidities. Identification of genetic variants contributing to metabolic variation will aid in understanding the pathophysiologic mechanisms underlying equine metabolic syndrome ultimately leading to the discovery of preventative measures and therapeutic targets in horses, in addition to potentially shedding light on the molecular basis of metabolic trait variation in other mammalian species.

5.3 Methods

5.3.1 Phenotyping

The Morgan horse data consists of 286 individuals sampled from 54 herds located throughout the United States and a single Canadian herd. Eleven phenotypes were tested for SNP associations and included morphometric measurements: neck circumference to height ratio (NH) and girth to height ratio (GH); and blood concentration measurements: fasting glucose (GLU), fasting insulin (INS), glucose 75 minutes post oral sugar administration (GLU OST), insulin 75 minutes post oral sugar administration (INS OST), triglycerides (TG), non-esterified fatty acids (NEFA), adrenocorticotropin hormone (ACTH), leptin (LEP), and adiponectin (APN). A more detailed description of the phenotypic measurements can be found in Section 2.3.3. INS, TG, NEFA, ACTH,

LEP, and APN exhibited a skewed distribution and required log or square root transformations to achieve a more normal distribution. Phenotypes were scaled to zero mean and standard deviation equal to one.

5.3.2 Genotyping

The horses were genotyped on on the Illumina SNP50 array and imputed up to ~2 million genotypes from a cosmopolitan reference population comprised of 384 individuals representing 30+ breeds, including 43 horses from the Morgan breed using Beagle 4.0 software[211,212]). Genotypes were filtered with MAF $>$ **0.05**, and Hardy-Weinberg disequilibrium p-values $>$ **0.00001** yielding 877,608 SNPs.

5.3.3 Polygenic and shared environment variance estimates

The polygenic variance (a.k.a."chip heritability") and shared environment variance estimates were determined using the GEMMA software.[213] Polygenic variance estimates were determined from a mean centered genomic relationship matrix comprised of all 845,234 autosomal SNPs and adjusted for age and sex covariates. Shared environment estimates were determined from a pairwise relationship matrix where pairwise relationships between individuals sharing an environment were represented by 1 and pairwise relationships between individuals not sharing an environment were represented by zero. Shared environment estimates were also adjusted for age and sex covariates.

5.3.4 Association testing

The improved linear mixed model as described in Section 4.3.5 was used to determine phenotype-genotype associations. The improved LMM more accurately models the true genetic architecture by applying a Bayesian method to initially rank SNPs combined with stepwise feature selection to select an optimal subset of SNPs for inclusion in the GRM. Recently, Bayesian methods have been developed that apply non-infinitesimal model assumptions to produce improved genetic prediction accuracy.[207,228] The improved LMM extends this concept to association testing but still provides the traditional GWAS frequentist test statistic. Briefly, the improved LMM algorithm is a basic 3 step

process:

- **Step 1:** Bin each chromosome into 500 kilobase (KB) segments and select the most influential SNP in addition to the two immediately adjacent SNPs to represent each bin.

- **Step 2:** Perform stepwise feature (bin) selection to build the select SNP GRM.

- **Step 3:** Perform LMM GWAS with select SNP GRM in place of an all SNP GRM.

In Step 1, BSLMM was performed to rank SNPs based on model frequency (the proportion of iterations that the SNP was included in the subset of markers deemed to have a large effect). The SNP with the highest BSLMM model frequency in each 500KB bin was deemed the most influential and selected to represent the bin along with the two immediately adjacent SNPs. The GEMMA software[207] was used to perform the BSLMM. The phenotypes were first adjusted for covariates age and sex prior to performing the BSLMM analysis as the software does not incorporate inclusion of covariates. Covariate effects were estimated from a linear mixed effect model including a random intercept for the clusters. 550,000 MCMC iterations were performed (including 50,000 burn-in iterations).

Step 2 involves a stepwise feature selection model building process to determine the ideal bins for inclusion in the select SNP GRM. First, a null model which includes fixed effect covariates and a random intercept term for each genetic cluster (farm) is compared with an alternative model that includes an additional polygenic random effect determined from a select SNP GRM calculated from the SNPs representing the top ranked bin from the BSLMM. A likelihood ratio test is performed to determine if inclusion of SNPs representing the bin into the select SNP GRM significantly improve the model. The degrees of freedom are determined from the difference in the effective degrees of freedom of each model.[214,215] If inclusion of the bin is significant, the alternative model becomes the new null model and is tested against an alternative model that is the same as the new null model with the exception that the SNPs representing the next highest ranked bin are also included in the calculation of the select SNP GRM. Under the condition where the alternative model is not significantly improved over the null model, the

SNPs representing the current bin are removed from the select SNP GRM in exchange for the SNPs representing the next ranked bin. The process continues until all bins have been tested for inclusion in the select SNP GRM.

After completion of Step 2, the select SNP GRM is built and ready for inclusion in Step 3, the improved LMM GWAS. During the testing of a SNP in the final improved LMM GWAS, SNPs within 1MB of a test SNP are removed from the select SNP GRM to avoid proximal contamination.[206,216] The final GWAS model includes covariates and a random intercept for each genetic cluster in addition to the select SNP GRM. The FaST-LMM software is used to estimate parameters in steps 2 and 3 by maximizing the log-likelihood using the algorithms of Lippert et al.[206,217] and Widmer et al.[208] Bonferroni correction was used to set the genome-wide significance threshold (p-value$< \mathbf{5.70x10^{-8}}$), however Bonferroni corrections are considered conservative since Bonferroni assumes independence among SNPs and linkage disequilibrium violates this assumption. Therefore, a suggestive threshold was set at p-value$< \mathbf{1x10^{-5}}$, similar with previous equine genome-wide studies.[230–232]

5.3.5 Investigation of EMS candidate loci overlap with human metabolic syndrome loci

Human GWAS data were obtained from the PheGenI database (http://www.ncbi.nlm.nih.gov/gap/phegeni accessed on December 1 2015) using default search parameters. The genes nearest the top p-value for all EMS associated loci above the suggestive significance threshold were searched against the following human GWAS phenotype terms for 11 different metabolic syndrome components and co-morbities:

1. adiposity (*Abdominal Fat, Intra-Abdominal Fat*)
2. body mass (*Body Mass Index, Obesity, Waist Circumference, Waist-Hip Ratio, Body Weight Changes, Body Weights and Measures, Body Weight*)
3. diabetes (*Diabetes Mellitus, Type 1 Diabetes, Type 2 Diabetes, Diabetic Nephropathy, Diabetic Retinopathy*)
4. glucose (*Glucose, Glycosylated Hemoglobin A*)
5. insulin and insulin resistance

6. lipids (*Apolipoprotein A-I, Apolipoproteins C, Cholesterol, HDL, LDL, Lipoproteins, VLDL, Triglycerides*)

7. cardiovascular (*Arteries, Atrial Fibrillation, Blood Flow Velocity, Blood Pressure, Coronary Artery Disease, Echocardiography, Heart Failure, Heart Rate, Hematocrit, Hypertension, Myocardial Infarction, Blood Coagulation Factors, Left Ventricular Hypertrophy*)

8. stroke

9. adiponectin

10. C-Reactive Protein

11. monocytes (*Monocyte Chemoattractant Protein-1, Monocytes*)

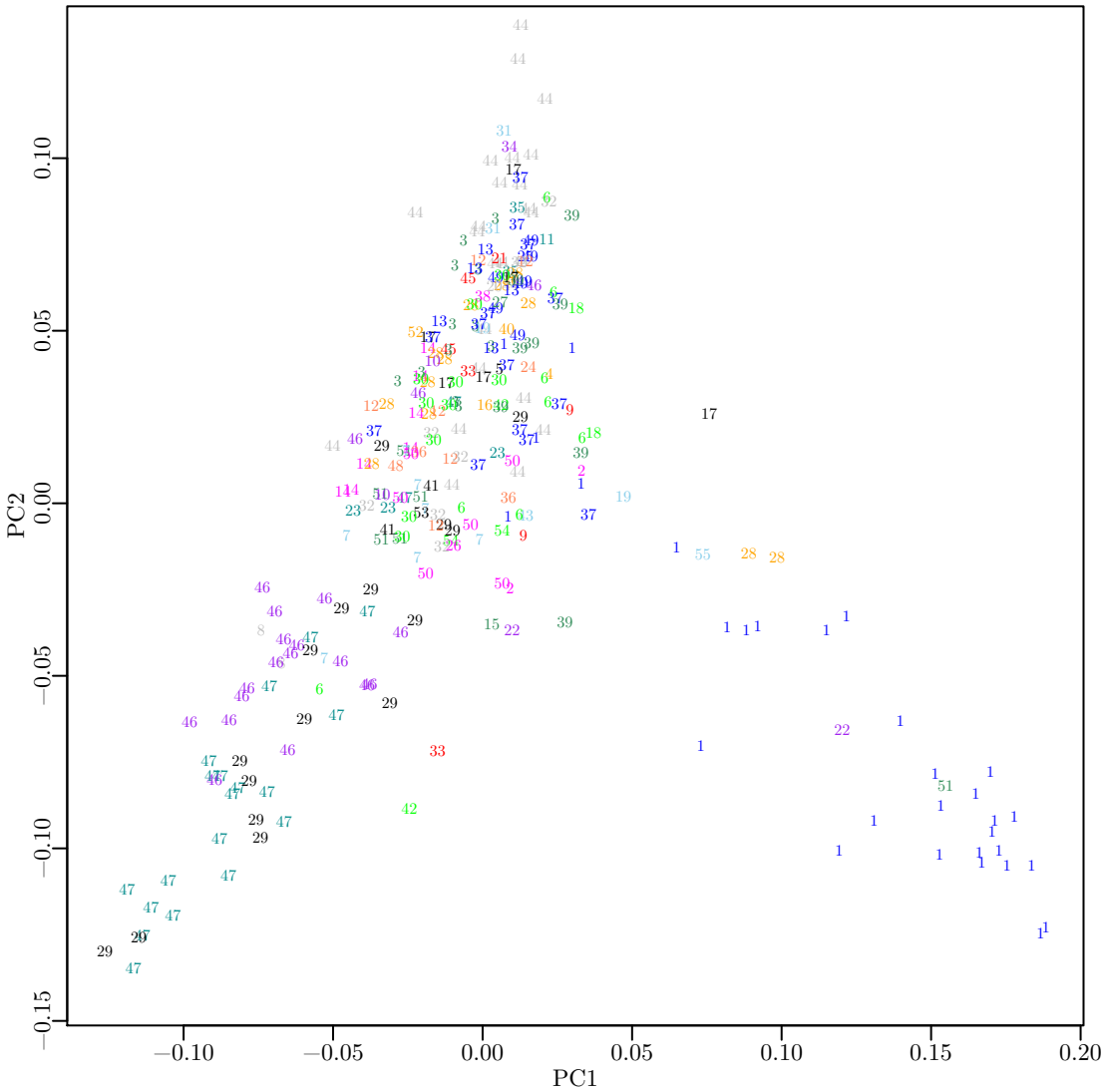
5.4 Results

5.4.1 Heritability estimation in a population featuring familial relationships and shared environments

Increased genetic similarity existed among individuals sampled form the same farm, illustrated in the PCA plot presented in Figure 5.1. Individuals sampled from the same farm are identified by the same number and share a similar color. It is easy to see the non-random allocation of numbers and colors, indicating individuals sampled from the same farm are more genetically similar than two individuals sampled from different farms.

Metabolic traits were also more similar among individuals sampled from the same farm. Figure 5.2 illustrates clustering of metabolic trait values near the farm median value. A linear relationship should not be easily discernible between individual trait values and the farm median in the case of independent observations. In Table 5.1, the "Farm ICC" column quantifies the increased phenotype similarity of individuals sampled from the same farm in comparison to individuals sampled from different farms. The farm intracluster correlation coefficient (ICC) is interpreted as 1) the proportion of variance at the farm level and 2) the correlation between measured outcomes for two

Figure 5.1: PCA illustration of increased genetic similarity at the farm level, each farm is represented by a different numeric value. Horses sampled from the same environment are identified with the same color (note: colors are repeated for the 55 herds).



randomly drawn individuals from the same cluster (farm), see (Section 1.3) for review of ICC.

Figure 5.2: Illustration of metabolic trait clustering in shared environments: neck circumference:height ratio(a), girth:height ratio(b), fasting glucose(c), fasting insulin(d), post-OST glucose(e) and post-OST insulin(f). Trait values centered to zero mean and scaled to one standard deviation variance.

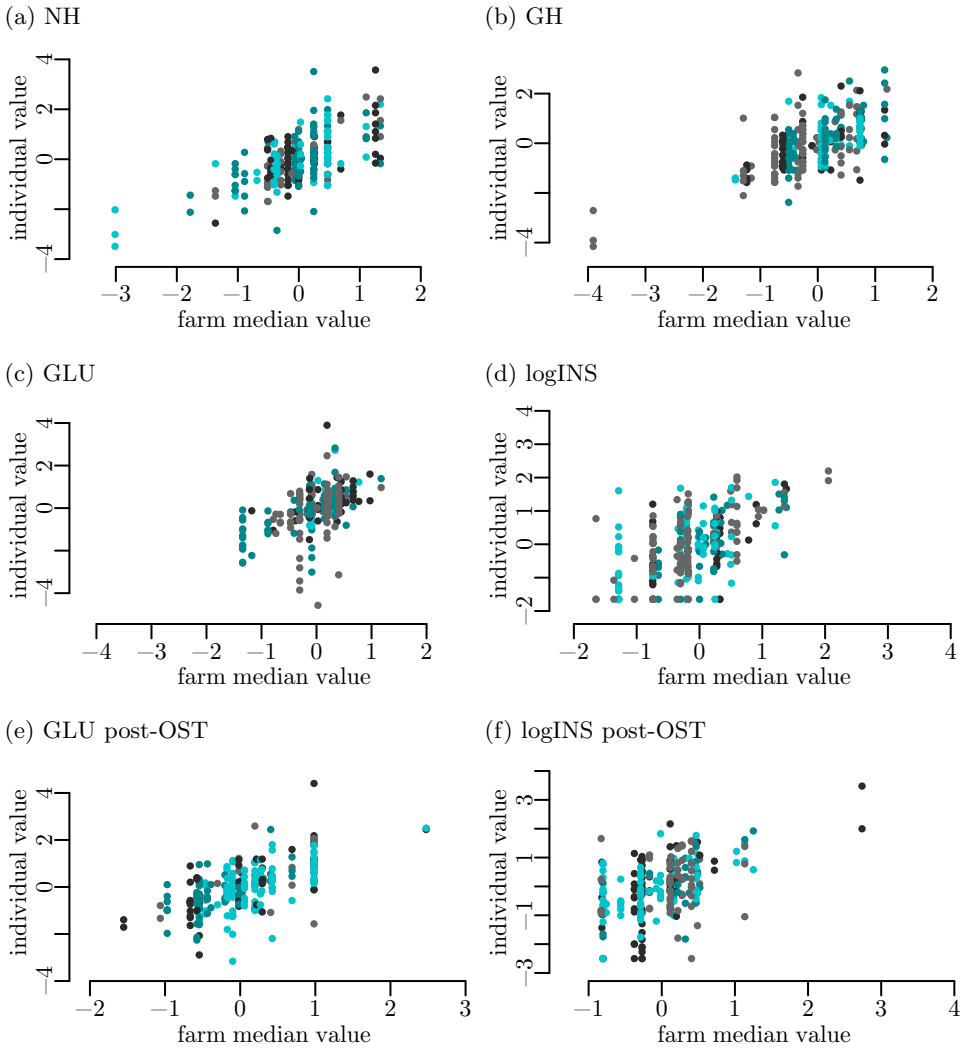
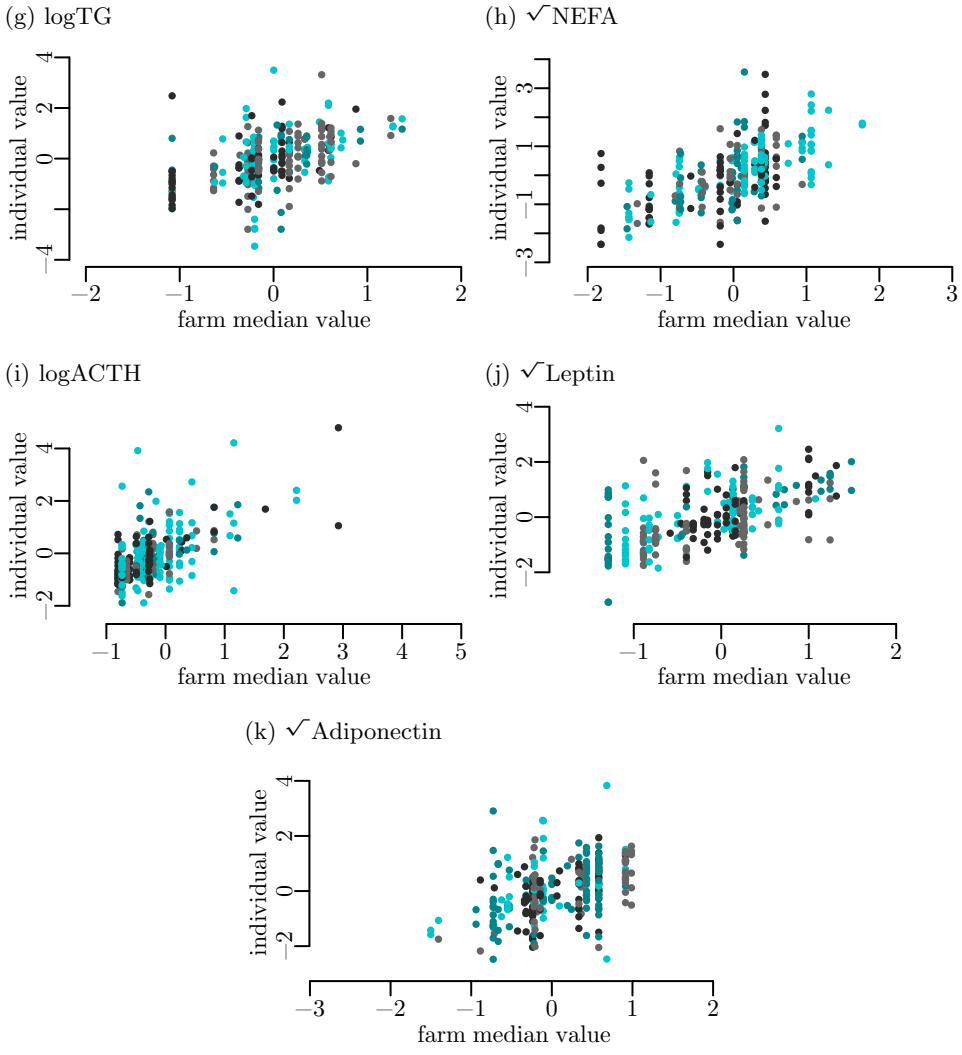


Figure 5.2 *Continued from previous page*: Illustration of metabolic trait clustering in shared environments: triglycerides(g), NEFA(h), ACTH(i), leptin(j), and adiponectin(k). Trait values centered to zero mean and scaled to one standard deviation variance.



Polygenic variance estimates ("narrow sense chip heritability") are presented in Table 5.1. The polygenic variance estimate is an estimate of the amount of trait variance attributable to shared genetic backgrounds ("heritability") and is determined from a SNP based genomic relationship matrix, i.e. the degree to which individuals with increased genotype sharing are also more phenotypically similar.

An accurate estimate of heritability is not possible due to confounding of environmental factors correlated with the population structure in this dataset. Polygenic variance and environmental variance (farm ICC) are not mutually exclusive. Individuals sampled from the same farm may be more phenotypically similar due to both shared environmental factors and shared genetic background. Therefore, it is important to note that this estimate of heritability may be upwardly biased if environmental factors influence the trait and individuals with increased genotype similarity also share the same environment. Polygenic variance estimates ("heritability") obtained from a model that also simultaneously partitions the farm level variance are presented in Table 5.2.

It is important to note heritability estimates adjusted for farm effects will be conservative given increased genetic similarity of individuals sampled from the same farm. Variance decomposition of fasting glucose illustrates the confounding effect of shared environment and genetic background. The genetic background (polygenic variance=0.22) and environment level variance (farm ICC=0.17) for fasting glucose are fairly similar when determined separately in Table 5.1. However, in a model that simultaneously estimates the polygenic variance and farm environment variance (Table 5.2) one can easily see that the two variance estimates are correlated as the polygenic variance estimate goes to zero and the farm level variance remains at 0.18. On the other hand, the polygenic variance is relatively high for adiponectin ("heritability"=0.70) and the farm environment variance is relatively low (0.12) when estimated separately in Table 5.1. Likewise, in a model that simultaneously estimates the polygenic variance and farm environment variance (Table 5.2) for adiponectin, the polygenic variance remains relatively high ("heritability"=0.64) since the farm level variance was not large.

Table 5.1: Genetic and environmental variance estimates.

Trait	N=	Polygenic variance estimate a.k.a "chip heritability" (standard error)	Farm ICC (standard error)
NH	285	0.59(0.15)	0.44(0.08)
GH	285	0.55(0.18)	0.43(0.08)
GLU	284	0.22(0.12)	0.17(0.07)
INS	285	0.62(0.15)	0.39 (0.09)
GLU OST	268	0.54(0.15)	0.29(0.08)
INS OST	268	0.38(0.15)	0.23(0.08)
TG	284	0.29(0.14)	0.17(0.06)
NEFA	286	0.65(0.18)	0.31(0.07)
ACTH	284	0.49(0.20)	0.41(0.09)
LEP	280	0.54(0.13)	0.26(0.07)
APN	285	0.70(0.17)	0.12(0.05)

ICC=Intracluster correlation coefficient is a measure of homogeneity of sampled clusters that can be interpreted two ways: 1) proportion of total variance of an outcome accounted for by the clusters or 2) the correlation between measured outcomes for two randomly drawn individuals in the cluster.

Table 5.2: Polygenic variance estimate ("heritability") adjusted for farm effects.

	polygenic variance	farm variance	residual variance
NH	0.22(0.11)	0.43(0.08)	0.34(0.11)
GH	0.20(0.12)	0.44(0.08)	0.36(0.12)
GLU	0.00(0.00)	0.18(0.07)	0.82(0.07)
INS	0.25(0.12)	0.39(0.09)	0.36(0.11)
GLU OST	0.13(0.13)	0.30(0.08)	0.57(0.14)
INS OST	0.20(0.13)	0.25(0.08)	0.55(0.13)
TG	0.00(0.00)	0.18(0.06)	0.82(0.06)
NEFA	0.00(0.00)	0.32(0.08)	0.68(0.08)
ACTH	0.25(0.11)	0.44(0.08)	0.30(0.11)
LEP	0.18(0.13)	0.24(0.08)	0.58(0.13)
APN	0.64(0.16)	0.05(0.05)	0.31(0.15)

Note: polygenic variance estimates ("heritability") adjusted for farm effects is a conservative estimate of heritability given increased genetic similarity of horses sampled from the same farm.

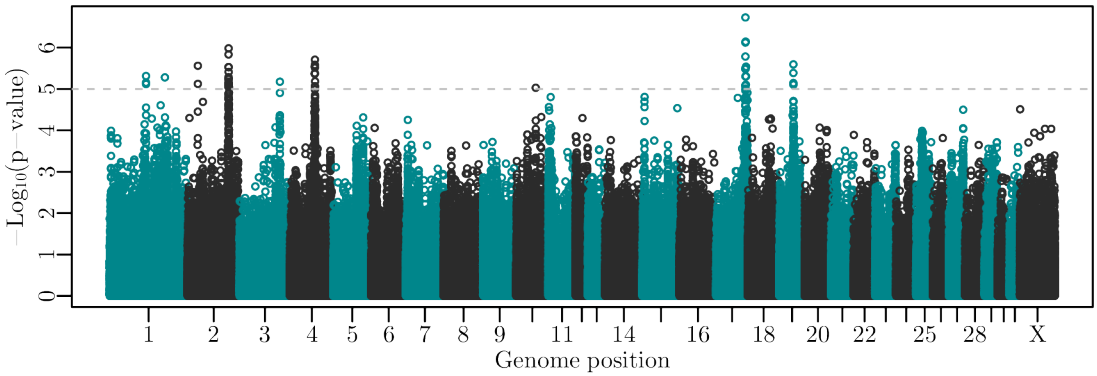
5.4.2 GWAS results

GWAS was performed on 286 Morgan horses from 55 different herds for 11 metabolic traits. The improved linear mixed model (LMM) described in Chapter 4 was utilized to test for associations while controlling for shared environment and polygenic background. Regional plots of candidate loci for all traits featuring color-coded $\boldsymbol{r^2}$ correlation values with the top SNP in the region are illustrated in Appendix Figures D.1 to D.11. Results determined from a standard LMM are reported in Section D.2.

Measures of adiposity

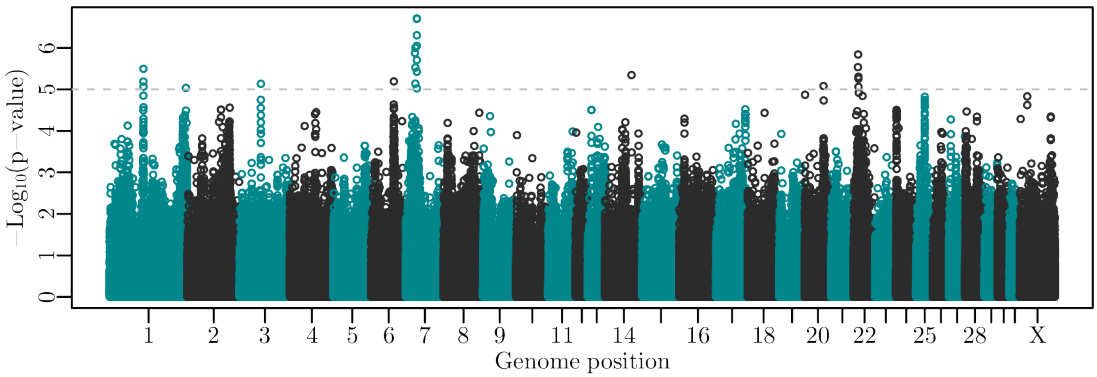
Manhattan plots for measures of regional adiposity (NH) and more generalized adiposity (GH), in addition to a listing of suggestive (p-value$<1\text{x}\mathbf{10^{-5}}$) and genome-wide (p-value$<\mathbf{5.70}\text{x}\mathbf{10^{-8}}$) significant loci are presented in Figures 5.3 and 5.4. Genome-wide significant variants were not detected for NH or GH, although several loci attained suggestive significance.

Figure 5.3: Manhattan plot of GWAS results for neck circumference to height ratio. Bonferonni threshold: p-value$<\mathbf{5.70x10^{-8}}$; suggestive threshold (dashed line): p-value$<\mathbf{1x10^{-5}}$. Loci reaching genome-wide and/or suggestive significance are listed below.



chromosome	position	alleles	maf	beta estimate	standard error	p-value	nearest gene
17	75543887	C\T	0.30	0.210	0.040	1.86E-07	*FAM155A*
2	93491935	A\G	0.36	0.201	0.041	1.04E-06	*PCDH18*
4	59620539	A\G	0.08	0.200	0.042	1.96E-06	*CREB5*
19	32821180	C\T	0.23	0.198	0.042	2.53E-06	*SENP5*
2	23825094	A\C	0.08	-0.194	0.041	2.74E-06	*ZNF362*
1	85359274	A\C	0.09	0.192	0.042	4.85E-06	*CCSER2*
1	128863092	C\T	0.06	-0.180	0.039	5.23E-06	*RAB8B*
3	96066755	G\A	0.08	-0.194	0.043	6.66E-06	*PCDH7*
10	52269721	C\T	0.11	-0.189	0.043	9.32E-06	*GRIK2*

Figure 5.4: Manhattan plot of GWAS results for girth to height ratio.
Bonferonni threshold: p-value$<\mathbf{5.70x10^{-8}}$; suggestive threshold (dashed line): p-value$<\mathbf{1x10^{-5}}$. Loci reaching genome-wide and/or suggestive significance are listed below.

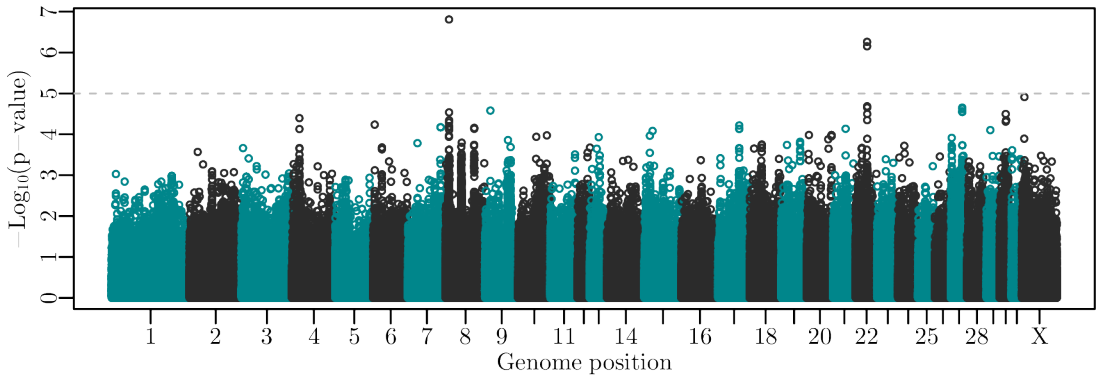


chromosome	position	alleles	maf	beta estimate	standard error	p-value	nearest gene
7	26154989	G\A	0.490	-0.198	0.038	1.95E-07	*MPZL2*
22	10321507	A\G	0.200	0.189	0.039	1.44E-06	*SEL1L2*
1	78254375	G\A	0.250	0.186	0.04	3.19E-06	*MAP10*
14	67885725	C\A	0.080	0.165	0.036	4.52E-06	*SLCO4C1*
6	50144883	A\C	0.440	-0.173	0.038	6.42E-06	*SOX5*
3	48758732	T\C	0.140	-0.184	0.041	7.34E-06	*GPRIN3*
20	46160950	A\G	0.280	-0.172	0.039	8.29E-06	*GPR115*
1	181616615	T\C	0.380	-0.174	0.039	9.21E-06	*MDGA2*

Glucose and insulin

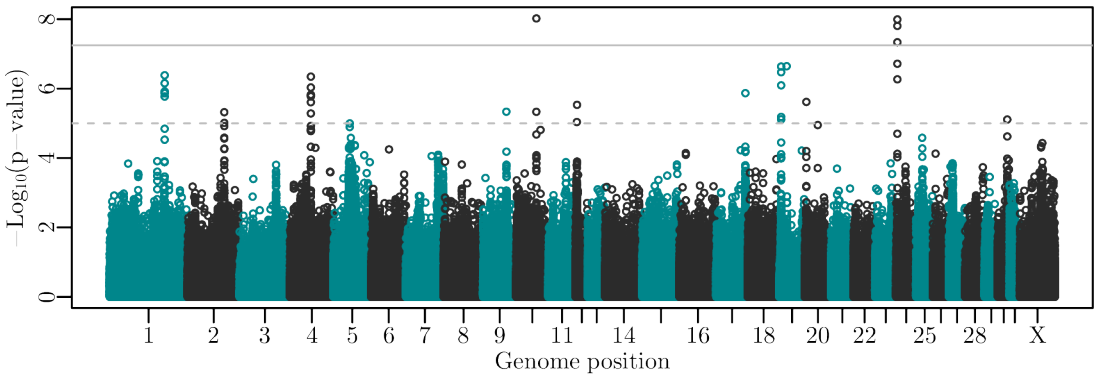
Manhattan plots and listing of suggestive and genome-wide significant candidate loci for both fasting and 75 minutes post-oral sugar challenge glucose and insulin levels are presented in Figures 5.5 to 5.9. Two genome-wide significant loci were detected for fasting insulin. The *GRIK2* locus on Equus caballus (ECA) autosome 10 illustrated in Figure 5.7 was the top reported fasting insulin locus in addition to reaching suggestive significance for neck circumference to height ratio (Figure D.1i). A second variant near *ATG14* on ECA 24 (see Figure D.4b) also reached genome-wide significance. Three loci presented in Figure D.6a-c located on ECA 21, 10, and 13 reached genome-wide significance for insulin levels 75 minutes following an oral sugar challenge. The variant on ECA 21 is near *ISL-1* (Figure 5.10) which encodes insulin gene enhancer protein. SNPs located within *FAM155A* on ECA 17 (Figures D.4g and D.1a) reached suggestive significance for both fasting insulin and NH.

Figure 5.5: Manhattan plot of GWAS results for fasting glucose.
Bonferonni threshold: p-value$<\mathbf{5.70x10^{-8}}$; suggestive threshold (dashed line): p-value$<\mathbf{1x10^{-5}}$. Loci reaching genome-wide and/or suggestive significance are listed below.



chromosome	position	alleles	maf	beta estimate	standard error	p-value	nearest gene
8	9312611	T\C	0.31	0.240	0.046	1.55E-07	*CRYBA4*
22	26963686	A\G	0.35	0.231	0.046	5.48E-07	*DLGAP4*

Figure 5.6: Manhattan plot of GWAS results for fasting insulin.
Bonferonni threshold (solid line): p-value$<\mathbf{5.70x10^{-8}}$; suggestive threshold (dashed line): p-value$<\mathbf{1x10^{-5}}$. Loci reaching genome-wide and/or suggestive significance are listed below.



chromosome	position	alleles	maf	beta estimate	standard error	p-value	nearest gene
10	53891179	T\C	0.15	0.192	0.034	9.54E-09	*GRIK2*
24	3597984	C\T	0.06	0.193	0.034	1.02E-08	*ATG14*
19	18465605	G\A	0.19	0.173	0.033	2.27E-07	*MFN1*
19	4295255	C\T	0.36	-0.177	0.034	2.32E-07	*OTOL1*
1	128192334	A\G	0.25	0.167	0.033	4.13E-07	*DAPK2*
4	49751990	A\G	0.48	0.176	0.035	4.53E-07	*AHR*
17	75473022	A\G	0.42	-0.163	0.034	1.35E-06	*FAM155A*
20	4333850	G\A	0.35	0.168	0.036	2.41E-06	*FAM50B*
12	2818242	A\G	0.21	0.164	0.035	2.94E-06	*COMMD9*
9	58656992	T\C	0.31	-0.157	0.034	4.63E-06	*TRPS1*
2	81787272	T\C	0.43	-0.149	0.033	4.79E-06	*GATB*
30	24253197	C\T	0.42	-0.150	0.034	7.78E-06	*KCNT2*

Figure 5.7: Regional plot of *GRIK2* locus association with fasting insulin. Color-coding represents correlation (r^2 value) with the reference SNP (dashed line)

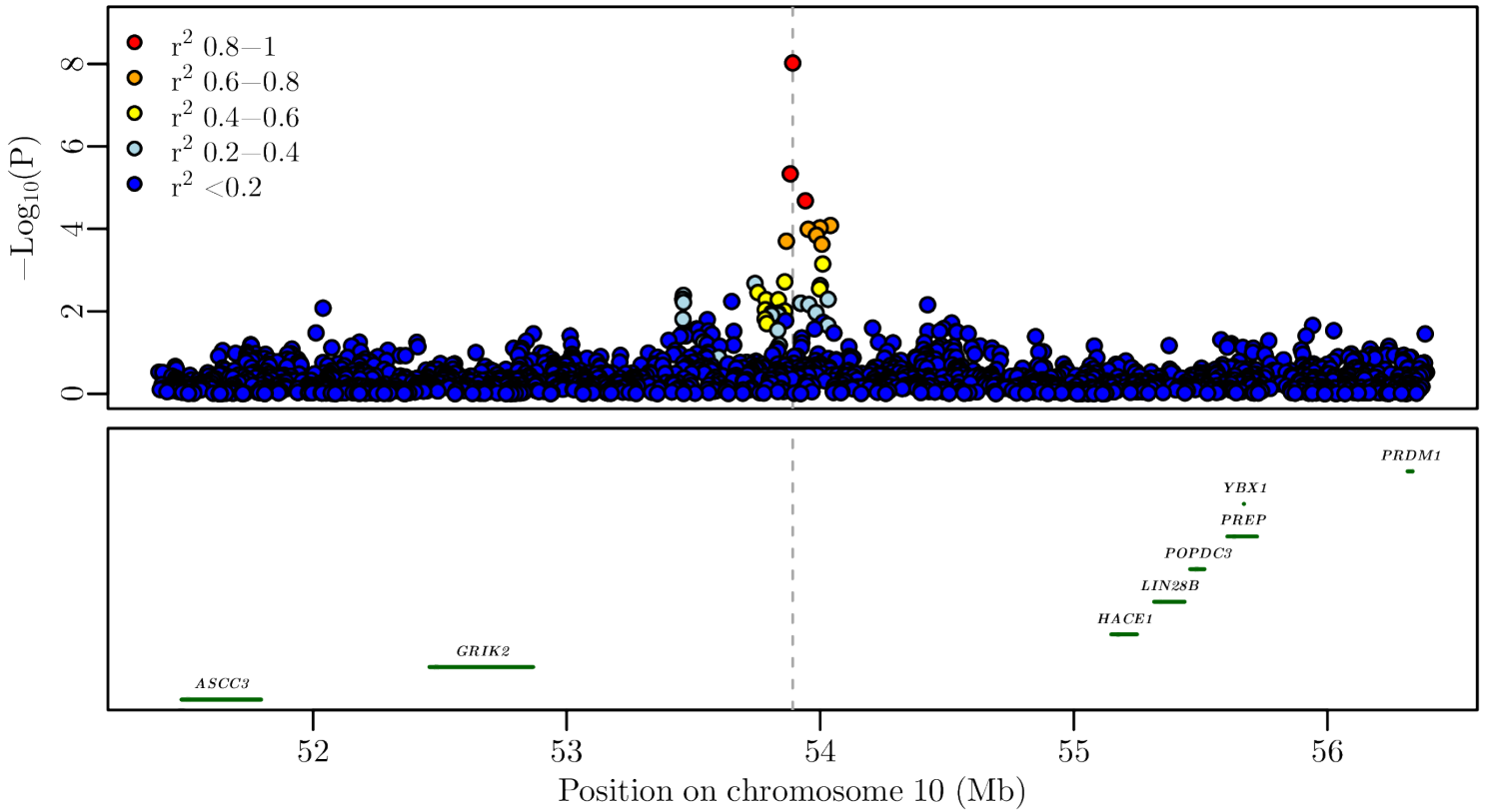
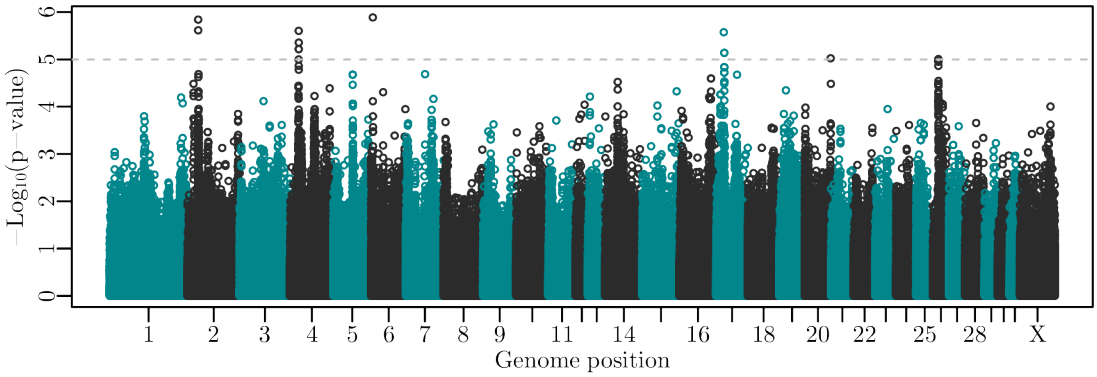
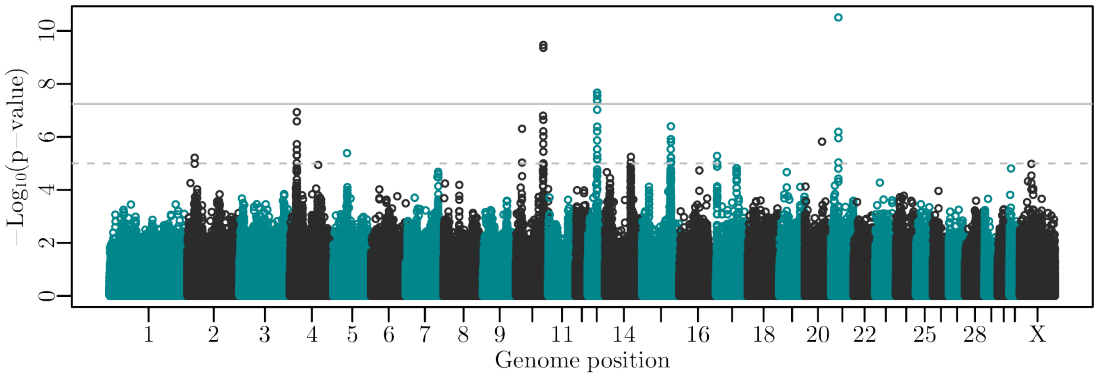


Figure 5.8: Manhattan plot of GWAS results for glucose 75 minute post-oral sugar test. Bonferonni threshold : p-value$<$**5.70x10^{-8}**; suggestive threshold (dashed line): p-value$<$**1x10^{-5}**. Loci reaching genome-wide and/or suggestive significance are listed below.



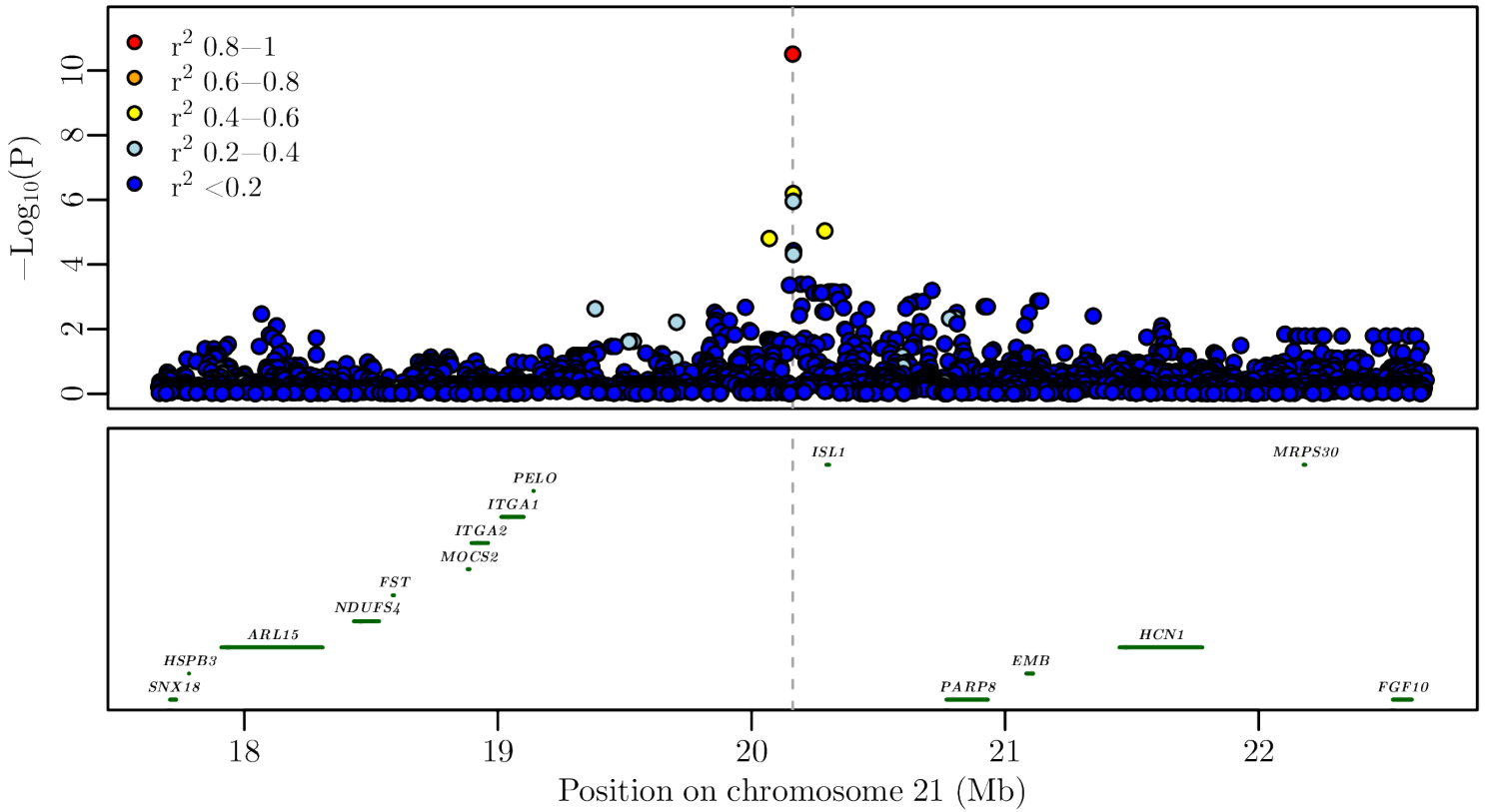
chromosome	position	alleles	maf	beta estimate	standard error	p-value	nearest gene
6	4059449	G\A	0.15	-0.211	0.044	1.30E-06	*SPAG16*
2	24516703	T\C	0.19	0.218	0.045	1.44E-06	*HDAC1*
4	21624394	T\C	0.22	-0.219	0.046	2.49E-06	*COBL*
17	20363782	T\G	0.05	-0.212	0.045	2.67E-06	*DLEU7*
20	63223615	T\C	0.32	0.230	0.052	9.47E-06	*RIMS1*
26	15146523	C\T	0.08	0.198	0.045	9.84E-06	*USP25*

Figure 5.9: Manhattan plot of GWAS results for insulin 75 minute post-oral sugar test. Bonferonni threshold (solid line): p-value$<\mathbf{5.70x10^{-8}}$; suggestive threshold (dashed line): p-value$<\mathbf{1x10^{-5}}$. Loci reaching genome-wide and/or suggestive significance are listed below.



chromosome	position	alleles	maf	beta estimate	standard error	p-value	nearest gene
21	20162158	C\A	0.12	-0.226	0.034	3.11E-11	*ISL1*
10	71766565	G\A	0.17	0.225	0.036	3.39E-10	*RNF217*
13	22385480	C\T	0.07	0.190	0.034	2.13E-08	*ZKSCAN2*
4	17084310	T\C	0.16	-0.175	0.033	1.16E-07	*TNS3*
15	71121201	T\C	0.42	-0.175	0.035	3.99E-07	*NCOA1*
10	16809864	A\C	0.34	-0.163	0.032	4.93E-07	*DACT3*
20	42730981	G\T	0.08	0.160	0.033	1.51E-06	*VEGFA*
5	39031828	C\T	0.18	0.154	0.034	4.10E-06	*OR10K1*
17	1746473	C\A	0.24	0.163	0.036	5.26E-06	*XPO4*
14	65684935	A\G	0.13	-0.155	0.034	5.67E-06	*NUDT12*
2	16703018	G\A	0.15	0.150	0.033	6.10E-06	*HIVEP3*

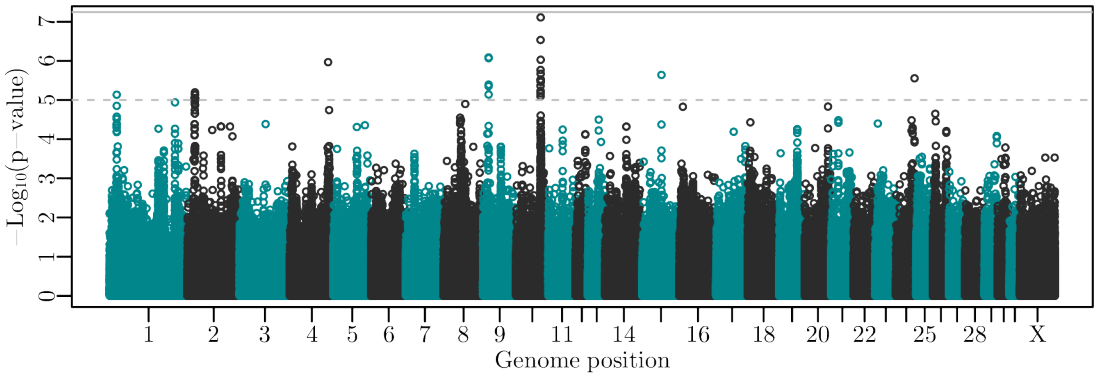
Figure 5.10: Regional plot of *ISL1* locus association with 75 minute post-oral sugar test insulin levels. Color-coding represents correlation (r^2 value) with the reference SNP (dashed line)



Lipids

Manhattan plots and listing of suggestive and genome-wide significant candidate loci for triglyceride and NEFA levels are presented in Figures 5.11 to 5.13. A variant within *FRK* on ECA 10 (Figure 5.12) attained near genome-wide significance (p-value $= \mathbf{7.8x10^{-08}}$) for triglyceride levels. Seven loci reached genome-wide significance for NEFA levels including variants in *ALK*, *NOVA1*, *FLVCR2*, *KLHDC1*, *SOX6*, and *ADAD1*, (see Figures D.8b-g). A SNP near *TSN* on ECA 18 (Figure 5.14) was the top ranked locus (p-value $= \mathbf{4.36x10^{-12}}$) for NEFA levels.

Figure 5.11: Manhattan plot of GWAS results for trigylcerides.
Bonferonni threshold (solid line): p-value$<\mathbf{5.70x10^{-8}}$; suggestive threshold (dashed line): p-value$<\mathbf{1x10^{-5}}$. Loci reaching genome-wide and/or suggestive significance are listed below.



chromosome	position	alleles	maf	beta estimate	standard error	p-value	nearest gene
10	64538148	T\C	0.15	0.207	0.038	7.80E-08	*FRK*
9	12293620	T\C	0.18	-0.190	0.038	8.18E-07	*JPH1*
4	95568021	T\C	0.17	0.179	0.037	1.08E-06	*TRBV19*
15	47489715	A\G	0.40	0.181	0.038	2.29E-06	*GPR75-ASB3*
24	43212532	A\C	0.28	-0.174	0.037	2.79E-06	*DIO3*
2	17237220	A\G	0.08	-0.183	0.040	6.35E-06	*CTPS1*
1	15523210	C\T	0.41	0.167	0.037	7.31E-06	*PNLIPRP1*

Figure 5.12: Regional plot of *FRK* locus association with triglycerides. Color-coding represents correlation (r^2 value) with the reference SNP (dashed line)

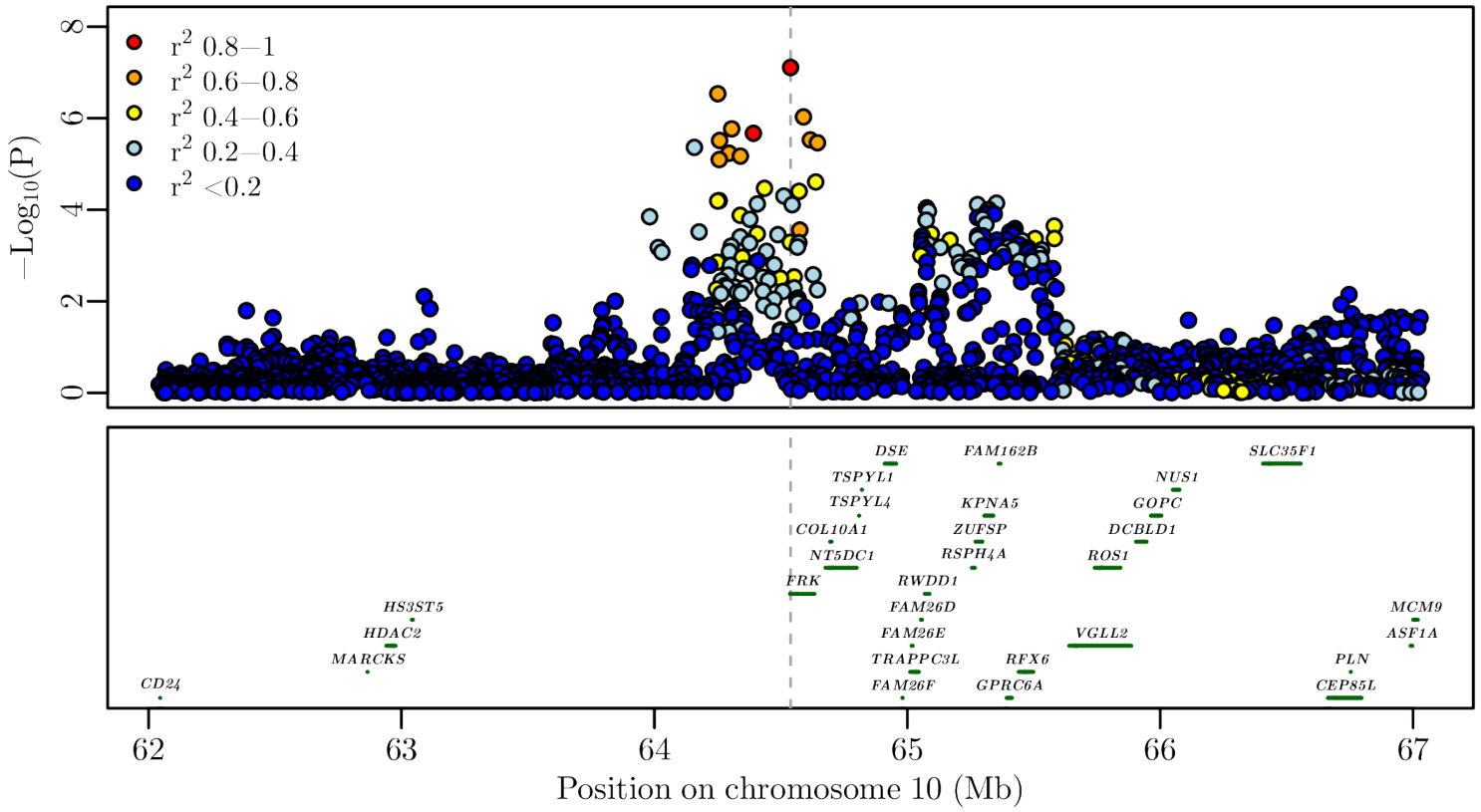
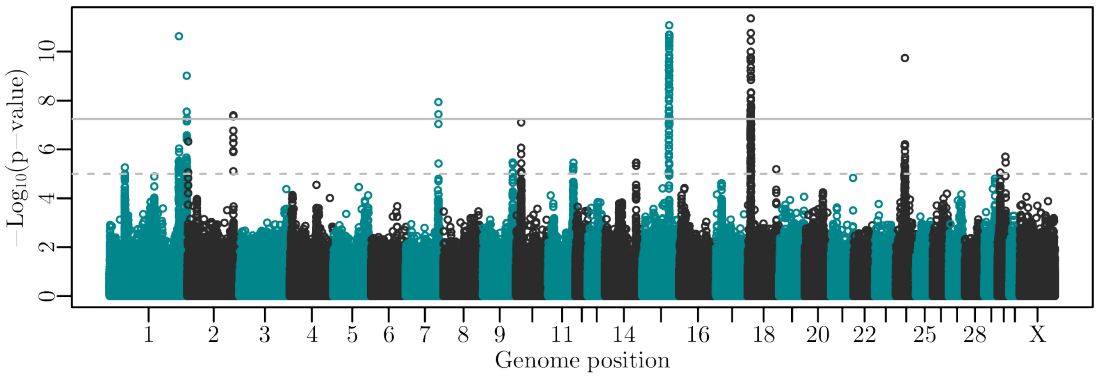
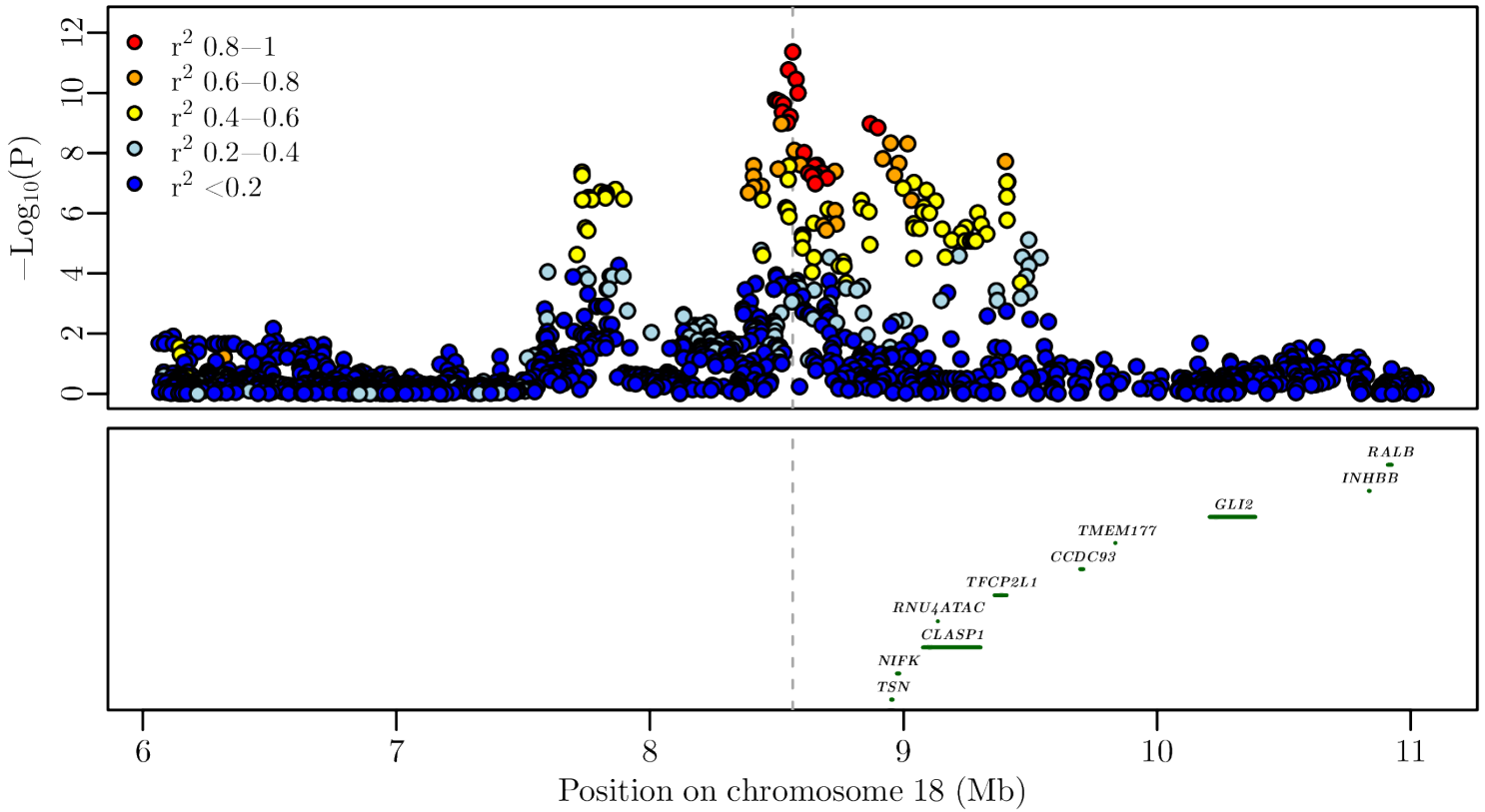


Figure 5.13: Manhattan plot of GWAS results for non-esterified fatty acids. Bonferonni threshold (solid line): p-value$<$**5.70x10^{-8}**; suggestive threshold (dashed line): p-value$<$**1x10^{-5}**. Loci reaching genome-wide and/or suggestive significance are listed below.



chromosome	position	alleles	maf	beta estimate	standard error	p-value	nearest gene
18	8562608	G\A	0.33	0.227	0.033	4.36E-12	*TSN*
15	66792714	G\T	0.11	0.212	0.031	8.33E-12	*ALK*
1	164702224	A\G	0.18	0.206	0.031	2.35E-11	*NOVA1*
24	20975408	G\T	0.35	0.200	0.031	1.83E-10	*FLVCR2*
1	183728003	G\A	0.23	-0.192	0.031	9.68E-10	*KLHDC1*
7	84765240	T\C	0.46	0.180	0.032	1.16E-08	*SOX6*
2	105684803	G\T	0.11	0.165	0.030	4.03E-08	*ADAD1*
10	13860272	G\A	0.33	0.163	0.030	7.87E-08	*LYPD4*
2	1779499	T\C	0.08	0.155	0.031	4.75E-07	*OMA1*
30	19352250	T\G	0.09	0.145	0.031	1.98E-06	*BRINP3*
9	74443615	C\T	0.22	-0.139	0.030	3.40E-06	*ST3GAL1*
11	56669565	A\G	0.21	-0.145	0.031	3.50E-06	*HS3ST3B1*
14	80043766	C\T	0.09	0.146	0.032	3.54E-06	*RN7SKP34*
1	35305763	C\T	0.29	0.144	0.032	5.46E-06	*LGI1*
18	75270775	G\A	0.14	-0.136	0.030	6.47E-06	*KCTD18*
30	5917999	G\A	0.44	-0.136	0.031	8.83E-06	*CNST*

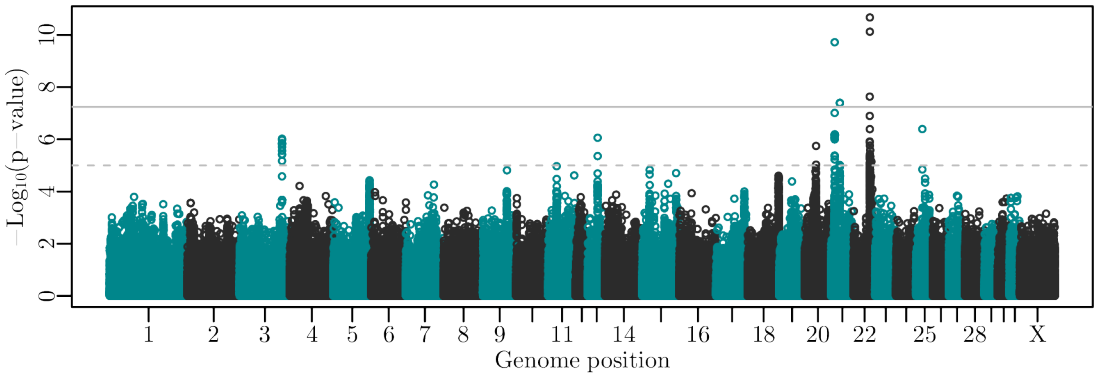
Figure 5.14: Regional plot of *TSN* locus association with NEFA. Color-coding represents correlation (r^2 value) with the reference SNP (dashed line)



ACTH

Manhattan plots and listing of suggestive and genome-wide significant candidate loci for ACTH levels are presented in Figure 5.15. Three loci reached genome-wide significance for ACTH levels, two of the loci contain candidate regions related to growth hormone. The nearest gene to the top ranked variant is *SNORD12* (Figure 5.16) which lies in a region identified to be associated with growth hormone deficiency in humans.[233] The third ranked locus contains a variant ~100KB from *GHR* on ECA 21 (Figure 5.17) that encodes the growth hormone receptor. *SEPP1* encodes selenoprotein P and is the nearest gene to the variant, however *GHR* is a more obvious biological candidate given growth hormone is also an anterior pituitary hormone and growth hormone receptor mutations impair growth hormone autofeedback signaling in pituitary tumors.[234]

Figure 5.15: Manhattan plot of GWAS results for adrenocorticotropin hormone. Bonferonni threshold (solid line): p-value$<\mathbf{5.70x10^{-8}}$; suggestive threshold (dashed line): p-value$<\mathbf{1x10^{-5}}$. Loci reaching genome-wide and/or suggestive significance are listed below.



chromosome	position	alleles	maf	beta estimate	standard error	p-value	nearest gene
22	37596454	C\T	0.23	0.219	0.033	2.14E-11	*SNORD12*
21	10999900	A\G	0.13	0.223	0.035	1.90E-10	*KIF2A*
21	23724278	T\C	0.18	-0.182	0.033	4.04E-08	*SEPP1*
25	14650611	T\C	0.09	0.164	0.032	4.05E-07	*EPB41L4B*
13	24087845	C\T	0.25	0.184	0.037	8.73E-07	*SCNN1B*
3	101439577	G\A	0.26	0.167	0.034	9.52E-07	*GBA3*
20	28706883	T\C	0.22	-0.152	0.032	1.79E-06	*OR2H1*

Figure 5.16: Regional plot of *SNORD12* locus association with ACTH. Color-coding represents correlation (r^2 value) with the reference SNP (dashed line)

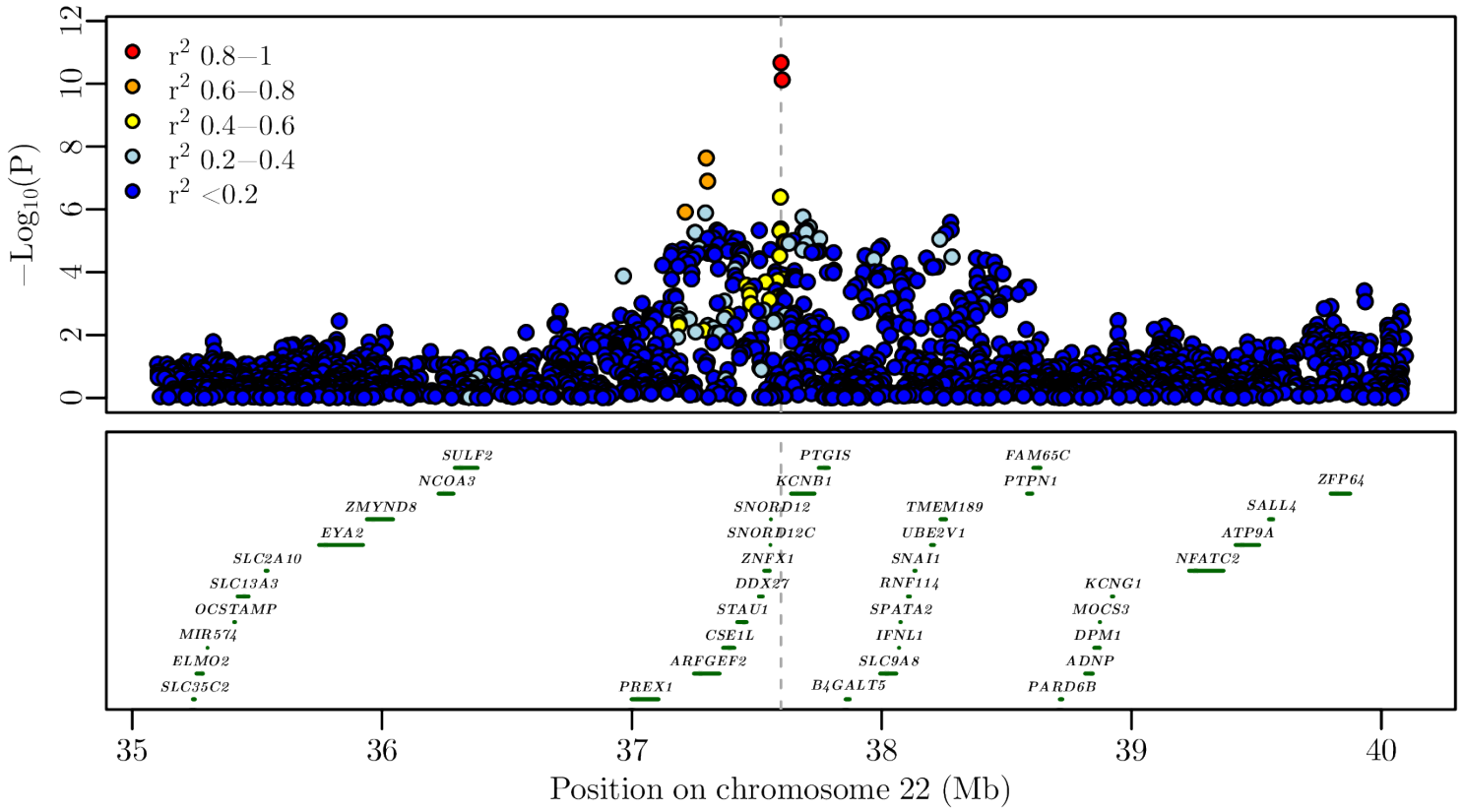
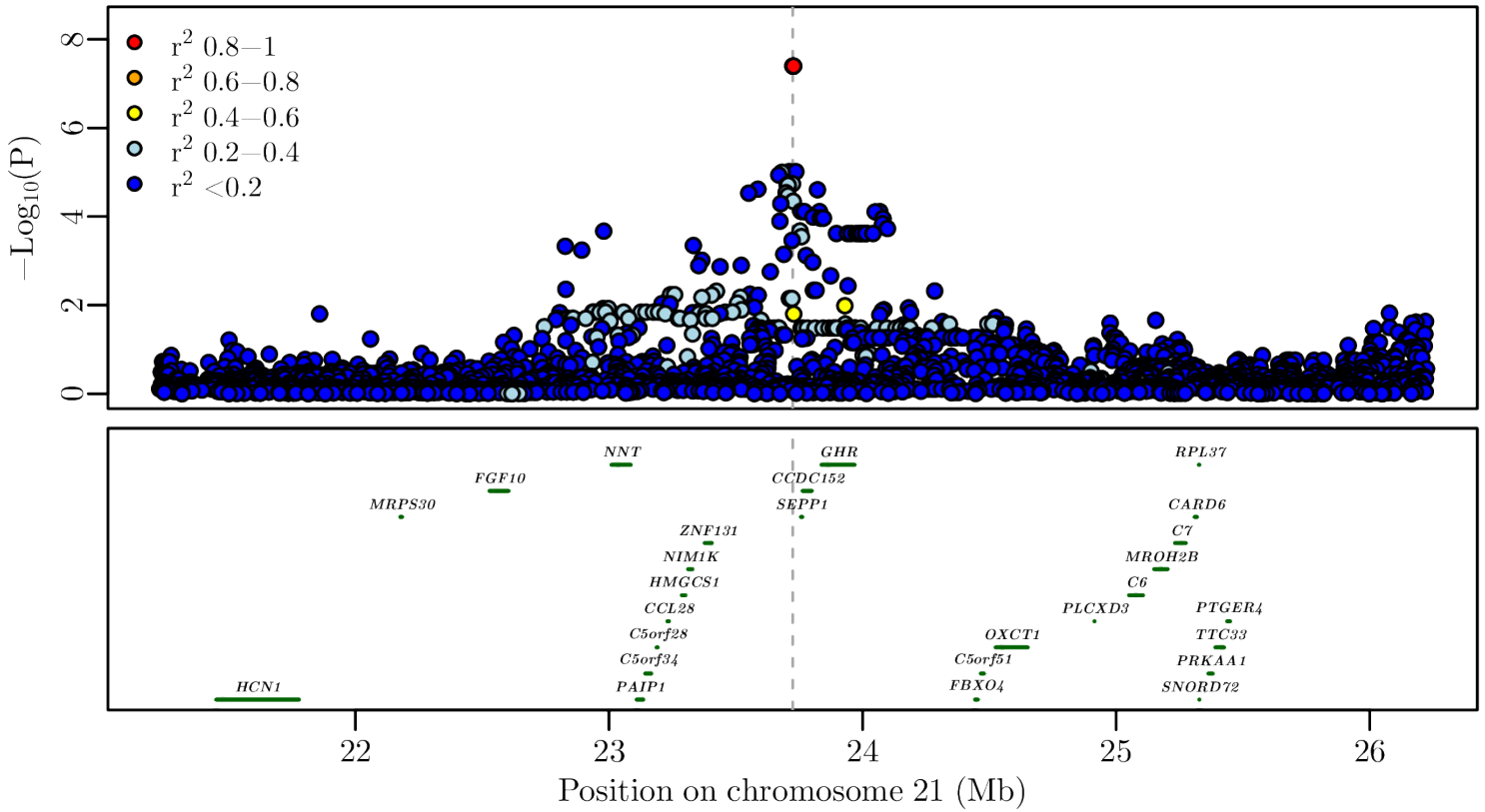


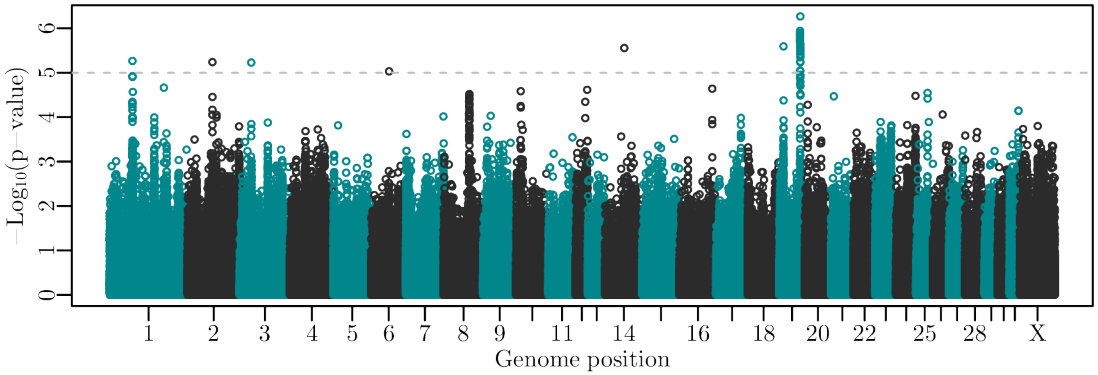
Figure 5.17: Regional plot of *GHR* locus association with ACTH. Color-coding represents correlation (r^2 value) with the reference SNP (dashed line)



Adipokines

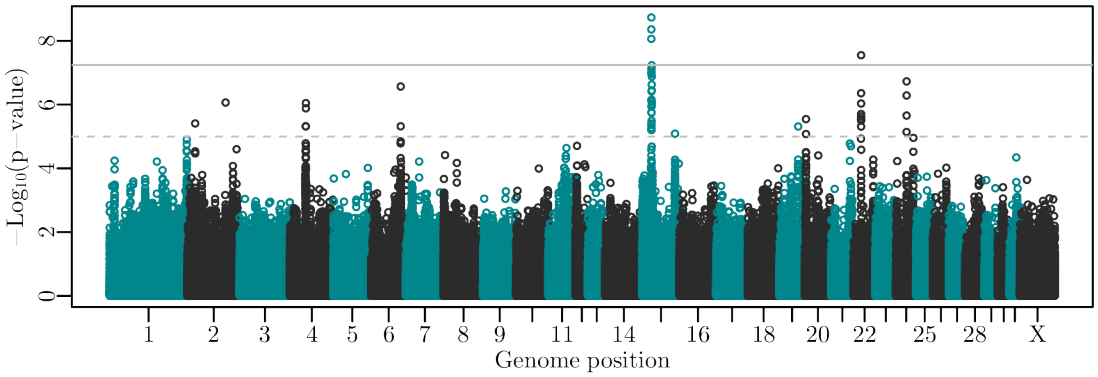
Manhattan plots and listing of suggestive and genome-wide significant candidate loci for leptin and adiponectin levels are presented in Figures 5.18 to 5.19. Genome-wide significant loci were not detected for leptin, although two loci on ECA 15 (Figure D.11a) and 22 (Figure D.11b) did reach genome-wide significance for adiponectin levels. In addition, a variant located within *NRXN3* (Figure 5.20) on ECA 24 attained near genome-wide significance (p-value $= \mathbf{1.88x10^{-07}}$) for adiponectin levels.

Figure 5.18: Manhattan plot of GWAS results for leptin.
Bonferonni threshold : p-value$<\mathbf{5.70x10^{-8}}$; suggestive threshold (dashed line): p-value$<\mathbf{1x10^{-5}}$. Loci reaching genome-wide and/or suggestive significance are listed below.



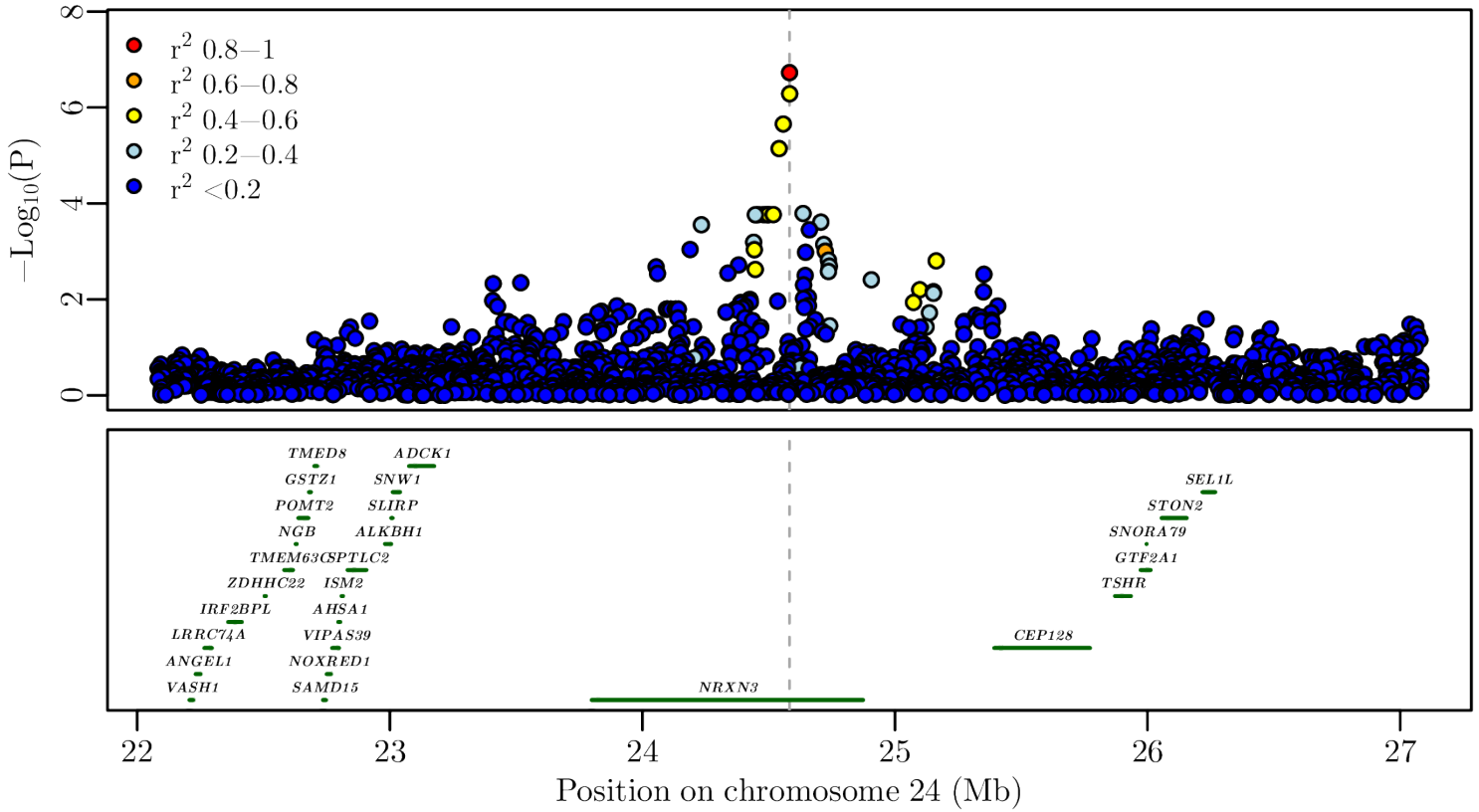
chromosome	position	alleles	maf	beta estimate	standard error	p-value	nearest gene
19	49581025	G\A	0.23	-0.187	0.037	5.42E-07	*CCDC54*
19	11111059	T\C	0.42	-0.180	0.038	2.56E-06	*RPL22L1*
14	47283907	G\A	0.23	0.181	0.039	2.79E-06	*MEGF10*
1	54004249	G\A	0.21	0.175	0.038	5.43E-06	*CTNNA3*
2	51826889	A\G	0.10	0.161	0.036	5.78E-06	*RHOBTB2*
3	25187849	C\T	0.18	-0.173	0.038	5.92E-06	*CNTNAP4*
6	38093791	G\A	0.13	0.155	0.035	9.32E-06	*KLRC1*

Figure 5.19: Manhattan plot of GWAS results for adiponectin.
Bonferonni threshold (solid line): p-value$<$**5.70x10^{-8}**; suggestive threshold (dashed line): p-value$<$**1x10^{-5}**. Loci reaching genome-wide and/or suggestive significance are listed below.



chromosome	position	alleles	maf	beta estimate	standard error	p-value	nearest gene
15	21018128	A\G	0.11	-0.200	0.033	1.85E-09	*ST6GAL-NAC2*
22	18132778	C\T	0.15	-0.194	0.035	2.83E-08	*SLC23A2*
24	24582154	G\A	0.15	-0.171	0.033	1.88E-07	*NRXN3*
6	68035083	G\A	0.11	0.178	0.035	2.71E-07	*DIP2B*
2	85771023	G\A	0.17	-0.169	0.034	8.62E-07	*TTC29*
4	37502360	C\T	0.23	-0.165	0.033	8.96E-07	*GNGT1*
20	3585516	G\T	0.35	-0.155	0.033	2.85E-06	*RIPK1*
2	17691609	T\C	0.11	-0.161	0.035	3.91E-06	*ZFP69B*
19	44097058	C\A	0.10	-0.154	0.034	4.86E-06	*GRAMD1C*
15	80933450	G\A	0.13	-0.151	0.034	8.17E-06	*TRIB2*

Figure 5.20: Regional plot of *NRXN3* locus association with adiponectin. Color-coding represents correlation (r^2 value) with the reference SNP (dashed line)



5.4.3 Control for false positives and prioritization of GWAS candidate genes

Q-Q plots of the observed versus expected test statistics presented in Figure 5.21 indicate absence of genomic inflation. In addition, overlap of candidate genes for human metabolic syndrome with candidate genes for equine metabolic syndrome (included all equine loci above the suggestive threshold) was investigated since population structure is unlikely to cause the same false positives in two different species and would provide strong evidence for a biological connection between the trait and the SNP in linkage disequilibrium with the candidate gene. Results of the human-equine metabolic syndrome candidate gene overlap investigation are reported in Table 5.3. The 93 identified equine metabolic syndrome candidate genes overlapped with 156 reported human gene metabolic trait associations (some genes associate with more than one human MetS trait). A similar investigation was performed on candidate genes identified using a standard linear mixed model GWAS and reported in Table D.12, the standard GWAS analysis also identified 93 candidate genes, however overlap with only 118 human gene metabolic trait associations were detected. In addition, 17 of the improved LMM GWAS reached genome-wide significance where as zero loci reached genome-wide significance using the standard LMM. 29 loci were identified by both the standard and improved LMM methods. The increased genome-wide significant findings and improved enrichment for human MetS candidate genes provided by using the improved LMM suggest the method is more powerful than a standard LMM in this study population.

Figure 5.21: Q-Q plots for EMS traits: neck circumference:height ratio(a), girth:height ratio(b), fasting glucose(c), fasting insulin(d), post-OST glucose(e) and post-OSTinsulin(f).

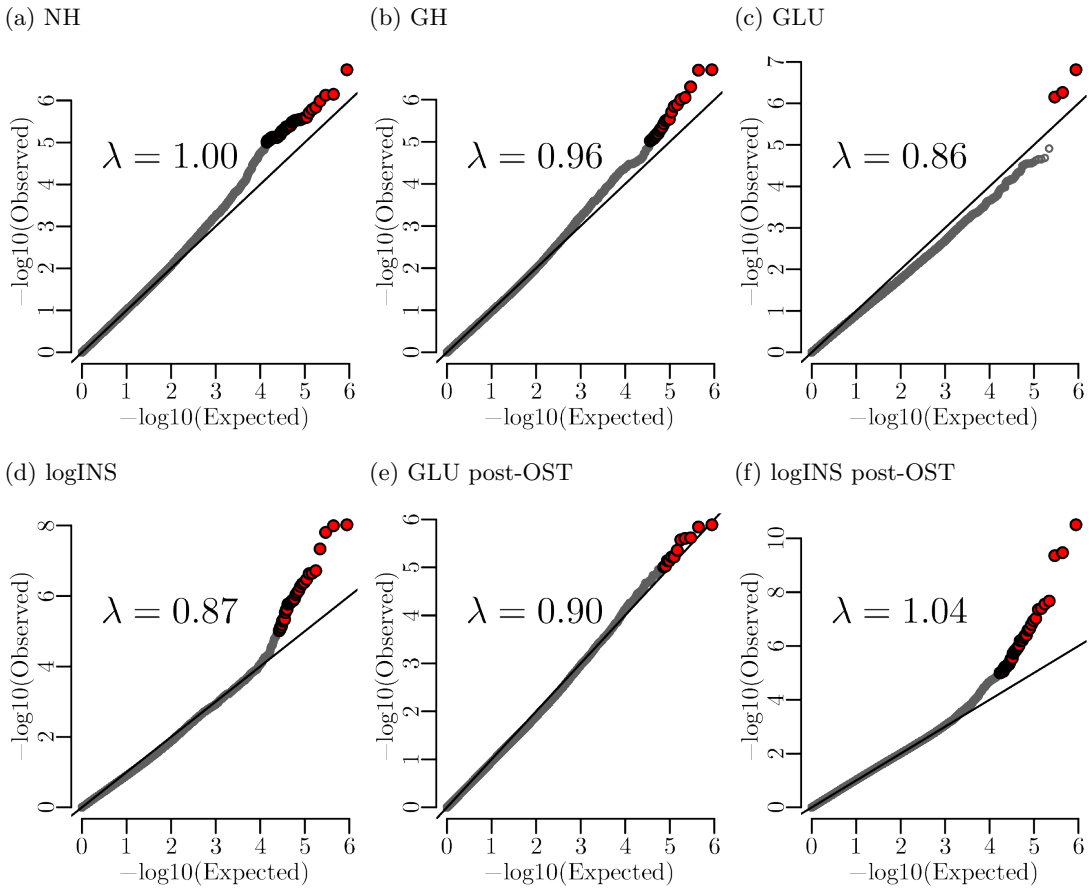


Figure 5.21 *Continued from previous page*: Q-Q plots for EMS traits: triglycerides(g), NEFA(h), ACTH(i), leptin(j), and adiponectin(k).

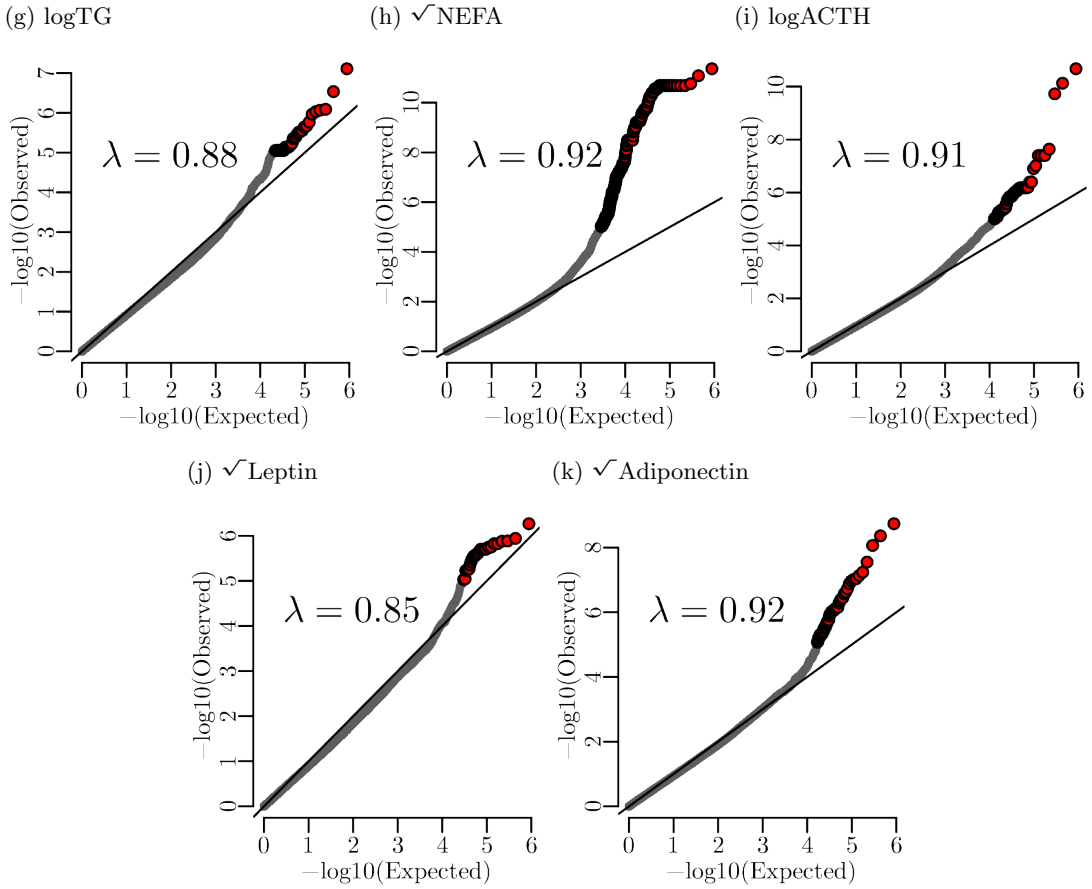


Table 5.3: Replication of human GWAS candidate genes for metabolic traits. Human GWAS data were obtained from the PheGenI database (http://www.ncbi.nlm.nih.gov/gap/phegeni accessed on December 1 2015) using default search parameters. The genes nearest the top p-value for all EMS associated loci above the suggestive significance threshold were searched against the following human GWAS phenotype terms for 11 different metabolic syndrome components and co-morbities: 1) adiposity (*Abdominal Fat, Intra-Abdominal Fat*), 2) body mass (*Body Mass Index, Obesity, Waist Circumference, Waist-Hip Ratio, Body Weight Changes, Body Weights and Measures, Body Weight*), 3) diabetes (*Diabetes Mellitus, Type 1 Diabetes, Type 2 Diabetes, Diabetic Nephropathy, Diabetic Retinopathy*), 4) glucose (*Glucose, Glycosylated Hemoglobin A*), 5) insulin and insulin resistance, 6) lipids (*Apolipoprotein A-I, Apolipoproteins C, Cholesterol, HDL, LDL, Lipoproteins, VLDL, Triglycerides*), 7) cardiovascular (*Arteries, Atrial Fibrillation, Blood Flow Velocity, Blood Pressure, Coronary Artery Disease, Echocardiography, Heart Failure, Heart Rate, Hematocrit, Hypertension, Myocardial Infarction, Blood Coagulation Factors, Left Ventricular Hypertrophy*), 8) stroke, 9) adiponectin, 10) C-Reactive Protein, and 11) monocytes (*Monocyte Chemoattractant Protein-1, Monocytes*).

			Human metabolic trait associations										
Gene	EMS trait	EMS p-value	*adiposity*	*body mass*	*diabetes*	*glucose*	*insulin/IR*	*lipids*	*cardiovascular*	*stroke*	*adiponectin*	*CRP*	*monocytes*
TSN	NEFA	4.4E-12		X				X	X	X			X
ALK	NEFA	8.3E-12		X	X	X	X	X	X				
SNORD12	ACTH	2.1E-11											
NOVA1	NEFA	2.4E-11		X				X	X				
ISL1	INS OST	3.1E-11						X	X				
FLVCR2	NEFA	1.8E-10		X				X					
KIF2A	ACTH	1.9E-10		X					X	X			
RNF217	INS OST	3.4E-10					X						
KLHDC1	NEFA	9.7E-10											
ST6GALNAC2	APN	1.9E-09											
GRIK2	INS	9.5E-09		X				X	X	X			
ATG14	INS	1.0E-08											
SOX6	NEFA	1.2E-08		X					X	X			
ZKSCAN2	INS OST	2.1E-08											

Table 5.3 *Continued on next page*

Table 5.3 *Continued from previous page*

Gene	EMS trait	EMS p-value	Human metabolic trait associations										
			adiposity	*body mass*	*diabetes*	*glucose*	*insulin/IR*	*lipids*	*cardiovascular*	*stroke*	*adiponectin*	*CRP*	*monocytes*
SLC23A2	APN	2.8E-08				X							
ADAD1	NEFA	4.0E-08			X								
SEPP1	ACTH	4.0E-08											
FRK	TG	7.8E-08		X	X			X	X	X			
LYPD4	NEFA	7.9E-08											
TNS3	INS OST	1.2E-07				X		X	X				
CRYBA4	GLU	1.6E-07											
FAM155A	NH	1.9E-07		X	X			X	X	X			
NRXN3	APN	1.9E-07	X	X	X	X		X	X				X
MPZL2	GH	2.0E-07											
MFN1	INS	2.3E-07											
OTOL1	INS	2.3E-07	X	X		X		X	X				
DIP2B	APN	2.7E-07											
NCOA1	INS OST	4.0E-07			X				X				
EPB41L4B	ACTH	4.1E-07											
DAPK2	INS	4.1E-07											
AHR	INS	4.5E-07						X	X				
OMA1	NEFA	4.8E-07						X					
DACT3	INS OST	4.9E-07											
CCDC54	LEP	5.4E-07		X				X	X				
DLGAP4	GLU	5.5E-07											
JPH1	TG	8.2E-07				X			X	X			
TTC29	APN	8.6E-07							X				X
SCNN1B	ACTH	8.7E-07											
GNGT1	APN	9.0E-07											
GBA3	ACTH	9.5E-07						X	X				
PCDH18	NH	1.0E-06		X					X				
TRBV19	TG	1.1E-06											

Table 5.3 *Continued on next page*

Table 5.3 *Continued from previous page*

Gene	EMS trait	EMS p-value	Human metabolic trait associations										
			adiposity	*body mass*	*diabetes*	*glucose*	*insulin/IR*	*lipids*	*cardiovascular*	*stroke*	*adiponectin*	*CRP*	*monocytes*
SPAG16	GLU OST	1.3E-06		X					X	X			
FAM155A	INS	1.4E-06		X	X			X	X	X			
SEL1L2	GH	1.4E-06											
HDAC1	GLU OST	1.4E-06											
VEGFA	INS OST	1.5E-06		X	X				X		X		X
OR2H1	ACTH	1.8E-06						X					
CREB5	NH	2.0E-06		X	X				X				
BRINP3	NEFA	2.0E-06		X		X		X	X	X			X
GPR75-ASB3	TG	2.3E-06											
FAM50B	INS	2.4E-06											
COBL	GLU OST	2.5E-06			X	X		X	X			X	
SENP5	NH	2.5E-06											
RPL22L1	LEP	2.6E-06							X	X			
DLEU7	GLU OST	2.7E-06						X	X				
ZNF362	NH	2.7E-06											
DIO3	TG	2.8E-06											
MEGF10	LEP	2.8E-06		X				X	X				
RIPK1	APN	2.9E-06											
COMMD9	INS	2.9E-06											
MAP10	GH	3.2E-06											
ST3GAL1	NEFA	3.4E-06		X				X	X	X			
HS3ST3B1	NEFA	3.5E-06							X				
RN7SKP34	NEFA	3.5E-06											
ZFP69B	APN	3.9E-06											
OR10K1	INS OST	4.1E-06											
SLCO4C1	GH	4.5E-06				X		X					
TRPS1	INS	4.6E-06		X	X	X	X	X	X	X			
GATB	INS	4.8E-06											

Table 5.3 *Continued on next page*

Table 5.3 *Continued from previous page*

Gene	EMS trait	EMS p-value	Human metabolic trait associations										
			adiposity	*body mass*	*diabetes*	*glucose*	*insulin/IR*	*lipids*	*cardiovascular*	*stroke*	*adiponectin*	*CRP*	*monocytes*
CCSER2	NH	4.9E-06											
GRAMD1C	APN	4.9E-06											
RAB8B	NH	5.2E-06						X					
XPO4	INS OST	5.3E-06							X				X
CTNNA3	LEP	5.4E-06											
LGI1	NEFA	5.5E-06							X				
NUDT12	INS OST	5.7E-06		X					X			X	
RHOBTB2	LEP	5.8E-06							X				
CNTNAP4	LEP	5.9E-06											
HIVEP3	INS OST	6.1E-06							X				
CTPS1	TG	6.4E-06											
SOX5	GH	6.4E-06		X	X			X	X				
KCTD18	NEFA	6.5E-06							X				
PCDH7	NH	6.7E-06	X	X	X			X	X	X			
PNLIPRP1	TG	7.3E-06											
GPRIN3	GH	7.3E-06							X				X
KCNT2	INS	7.8E-06		X				X	X	X	X	X	
TRIB2	APN	8.2E-06			X	X		X	X			X	
GPR115	GH	8.3E-06											
CNST	NEFA	8.8E-06											
MDGA2	GH	9.2E-06						X	X				
GRIK2	NH	9.3E-06		X				X	X	X			
KLRC1	LEP	9.3E-06											
RIMS1	GLU OST	9.5E-06						X	X				
USP25	GLU OST	9.8E-06				X			X				

5.5 Discussion

Extensive phenotyping combined with high density genotypes enabled localization of genetic loci associated with the pathophysiology of metabolic syndrome in horses. The results suggest substantial overlap of the biological processes underlying human and equine metabolic syndrome. Examples exist for monogenic traits where mutations in the same gene result in similar phenotypes in different species, such as mutations in the *MC1R* gene causing differences in hair color in humans[235] and variation in haircoat in horses and feathers in chickens.[236] Shared candidate genes have also been identified among humans, horses, and cattle for height, a classic polygenic trait.[220,222] Identification of shared candidate genes for metabolic syndrome in both humans and horses both provides validation for the gene as causal in each species since population stratification is unlikely to cause the same false positive in different species and secondly, potentially provides opportunity for exploring similar preventative and therapeutic management strategies.

Genome-wide and suggestive genetic variants were associated with measures of regional and generalized adiposity. The *FAM155* locus associated with neck circumference to height ratio (NH) in horses, in addition to equine fasting insulin levels. The *FAM155* locus has also demonstrated associations with human metabolic traits including: obesity in adults who received cranial radiation therapy as a child cancer patient,[237] C-reactive protein levels,[238] and suggestive associations with human body mass,[239] lipoproteins,[240] and hypertension[241] in the Framingham Heart Studies. Similar to the *FAM155* locus, the *GRIK2* locus was also associated with equine NH and fasting insulin levels, in addition to associations with human metabolic traits body mass,[239] cholesterol,[240] lipoproteins,[240] and hypertension[241] in the Framingham Heart Studies. The equine NH locus *PCDH7* is associated with human cholesterol levels[242] and body mass index.[239] *PCDH7* is also a paralog of *PCDH18* another protocadherin gene associated with equine NH, providing evidence for protocadherins playing a role in equine regional adiposity.

Several genome-wide and suggestive genetic variants were associated with glucose and insulin fasting and post-oral sugar test levels. A genetic variant near *AHR* which encodes the aryl hydrocarbon receptor, a ligand activated transcription factor known to bind natural plant flavanoids in addition to endocrine disrupting chemicals such as

polycyclic aromatic hydrocarbons and dioxins making it an interesting candidate gene to pursue given increased interest in the potential role of endocrine disrupting chemical in the pathophysiology of metabolic syndrome. The *AHR* locus is also associated with human caffeine consumption[243] and heart rate.[244] Equine fasting insulin levels were also associated with the *TRPS1* locus which has demonstrated association with human cholesterol levels.[245,246] Equine fasting insulin *KCNT2* locus is associated with human adiponectin levels.[247,248] The top ranked SNP for insulin levels following an oral sugar challenge in horses is *ISL1*, which encodes insulin enhancer protein that binds to the enhancer region of the insulin gene and may play an important role in regulating insulin gene expression. The *VEGFA* equine post-OST insulin locus is another interesting candidate gene given associations with human adiponectin levels[249] and waist-hip ratio.[250]

The top ranked SNP for triglycerides located within the *FRK* gene has been associated with obesity in the Korean population,[251] in addition to a reported association with cholesterol level.[245] Numerous variants located within genes were associated with equine NEFA levels and also happen to be associated with multiple human metabolic syndrome phenotypes, for example, the *ALK* locus is associated with human body mass and diabetes-related traits in the Framingham Heart Study[239,248] and *NOVA1* is associated with body weight and cholesterol levels.[239,240]

Genome-wide and suggestive genetic variants were associated with adipokines leptin and adiponectin. The equine adiponectin locus *NRXN3* is also associated with human body mass index,[252] obesity[253] and waist circumference.[254] Human associations also exist for the equine adiponectin locus *TRIB2* with cholesterol and apolipoprotein C levels.[240]

The current study is limited by a relatively small sample size for detecting variants of small effect contributing to a polygenic trait. The statistical algorithm incorporates Bayesian methodology and has demonstrated increased power to detect polygenic variants. However, the method may provide varying results depending on initial parameter settings, thus it is important not to assume all identified associations are causal without additional validation studies. Replication of candidate genes observed in humans provides support for identified equine loci being causal, however a statistical test for enrichment of human metabolic syndrome candidate genes was not performed. Performing a test of this nature and also investigating enrichment of metabolic syndrome

related pathways are recommended and would aid in prioritization of GWAS candidate genes for more in depth analysis.

An additional point of discussion is the use of only the nearest gene relative to the top ranked locus SNP to search for overlapping human MetS GWAS identified loci. It is well known that the top ranked SNP for a particular locus tags a haplotype that may contain multiple genes and the gene harboring the causal variant may not necessarily be the closest gene. The procedure used in this study to identify overlapping horse and human metabolic syndrome loci circumvents this issue to a degree. For a given gene list, the PheGenI database (http://www.ncbi.nlm.nih.gov/gap/phegeni) reports human genes nearest or containing SNPs identified significant by GWAS. Therefore the reported human gene may not be the gene nearest the top ranked SNP for a particular human locus, it is simply a gene within a human associated locus. For both human and horse loci, the nearest gene is a place to begin further investigations, especially if the gene function is known to be biologically relevant to the trait. However, one must be mindful to consider genes more distal to the tag SNP when performing enrichment or sequence analysis. Future analyses could specifically test for increased enrichment of the genes nearest tag SNPs compared to other genes in the locus to support further investigation (sequence analysis) of the nearest gene.

A final limitation of the current study is the performance of only a univariate response model. Many of the metabolic syndrome phenotypes are correlated which may in part be due to genetic correlation, for example in the current study the *GRIK2* and *FAM155A* loci were associated with both NH and fasting insulin. Incorporating a multivariate response model may increase power to detect additional loci, however the source code for the improved LMM algorithm presented in Chapter 4 has not yet been modified to include multivariate responses.

In summary, the current study identified several genome-wide and suggestive genetic variants associated with equine metabolic syndrome phenotypes suggesting the syndrome has a highly polygenic architecture similar to human metabolic syndrome. Furthermore, numerous overlapping candidate genes were observed for equine and human metabolic traits, indicating similar biological processes may underly metabolic syndrome pathophysiology in both species. Future work will require validation of equine candidate genes in a second population. Haplotypic association analysis with

high-density genotypes will aid in narrowing the associated regions and identifying all positional candidate genes to include in sequence analysis to identify the causal variants.

Chapter 6

Conclusions and future directions

Human metabolic syndrome (MetS) is a global health concern due to the increase in prevalence and association with increased risk of cardiovascular disease and type 2 diabetes (T2D). Horses suffer from a similar metabolic condition associated with increased risk of laminitis, a crippling and potentially life-threatening condition of the foot. Debate exists in both the human and veterinary fields regarding the etiology and pathogenesis of metabolic syndrome and the mechanisms linking the syndrome to its secondary consequences. In both fields, conflicting reports generated from different investigations of the etiology and pathogenesis of metabolic syndrome reflect the complex, multifactorial nature of the condition. It is possible these two similar conditions may share some common underlying biological processes and pathogenesis and knowledge from one species may be transferable to the other.

In order to develop effective preventative and therapeutic strategies for managing the secondary consequences of these syndromes, further research is warranted to determine the underlying mechanisms. The objectives of this thesis were to 1) quantify the variation in metabolic phenotypes across horse/pony breeds and the impact of individual and environmental factors, and to identify differences in metabolic phenotype in horses/ponies with obesity and/or a history of laminitis, 2) to examine response to an oral sugar challenge and incretin biology (DPP-IV activity; insulin secretory and GLP-1 responses; SNPs in GCG and DPP4) in Morgan horses and Welsh ponies, and 3) identify genetic loci associated with equine metabolic trait variation.

6.1 Characterization of the EMS phenotype

6.1.1 Benefits of a multilevel, multivariate analysis of metabolic syndrome

In Chapter 2 a multilevel, multivariate cross-sectional analysis of metabolic trait variation was performed in a large cohort (610) of horses and ponies sampled from 166 farms located throughout the US in addition to a single Canadian farm. Metabolic trait measurements were found to cluster at the farm level, i.e. horses sampled from the same farm were more likely to have similar metabolic trait measurements than horses sampled from different farms. The multilevel, multivariate analytic approach accounts for the sampling bias (nuisance factor) due to violating the assumption of independent observation but simultaneously facilitates a much deeper analysis, including quantifying the impact of both individual and farm related factors on metabolic trait variation and determining the degree of variation and correlation (co-variation) between equine metabolic traits at the farm and individual horse level. Traits that exhibit a more pronounced correlation at the farm level compared to the individual horse level indicate factors shared at the farm level (e.g. sampling time of year, pasture composition) may also explain co-variation in metabolic traits.

6.1.2 Regional vs generalized adiposity correlate with different biochemical profiles

Analysis of trait co-variation revealed triglyceride (TG), insulin (INS), and adiponectin (APN) levels correlated more strongly with a measure of regional adiposity (neck circumference to height ratio [NH]) than a measure of generalized adiposity (girth to height ratio [GH]). Previous data suggest the nuchal ligament adipose tissue is more likely than other adipose depots to display an inflammatory phenotype in the horse and therefore may play a unique role in the pathogenesis of metabolic dysregulation,[133] data from Chapter 2 support this hypothesis. Interestingly, in humans, subcutaneous fat in the neck region has been found to be strongly associated with insulin resistance suggesting the possibility of shared underlying mechanisms.[134]

6.1.3 Equine metabolic traits vary with obesity and laminitis status

Adipokine measurements, leptin (LEP) and APN, exhibited divergent patterns of variation according to obesity and prior laminitis status. LEP levels were elevated in obese horses to a similar degree with and without a prior history of laminitis whereas APN levels were decreased in horses with a prior history of laminitis to a similar degree in non-obese and obese horses. In addition, LEP correlated more strongly with generalized adiposity than regional adiposity. These findings suggest LEP may be a good indicator of body fat mass in horses and APN may indicate the presence of "unhealthy" adipose. Horses with a history of laminitis exhibiting significant differences in regional adiposity, insulin, triglycerides, and adiponectin levels independent of obesity status, providing further support for these components as laminitis risk factors and indicating obesity is not an essential component affecting variation in these metabolic traits. The effect of obesity on INS and LEP levels may limit the utility of these biochemical markers as a diagnostic test for metabolic disturbances due to causes other than over nutrition. In our population, TG and APN levels were less affected by obesity and may have an increased utility for diagnosing metabolic disturbances resulting from causes other than over nutrition.

6.1.4 Equine metabolic traits vary with breed, gender, age, season, and diet

Breed differences were observed for several traits, for example, the Quarter Horse (QH) breed was the most divergent with lower INS/INS OST, TG, and LEP. Increased muscle mass in the QH breed may explain the observation of a more insulin sensitive phenotype. Breed differences provide support for genetic variation as a source of metabolic trait variation. However, breed differences for metabolic traits may pose limitations on the use of a single reference range for all breeds when utilizing metabolic traits as a diagnostic.

Gender differences were observed for several traits with mares having a significantly lower measure of regional adiposity (NH) and stallions having a lower measure of generalized adiposity (GH), suggesting gender differences in equine body fat distribution.

Mares also had higher TG and LEP levels. The gender differences observed in horses share similarity with human gender differences. Triglyceride levels have been reported to correlate with fat deposition sex differences in humans[146] and gender differences in LEP levels in humans are accounted for by percent body fat.[147]Similar to previous reports,[144,148] age was positively associated with INS and to a larger degree with ACTH levels.

Seasonal variation in traits was similar to previous reports, although the current study is limited by the fact that samples were obtained from each horse on only a single day of the year. Elevated ACTH in the fall months has been previously reported in horses.[149,150] However, in our study positive effects of latitude on ACTH levels were observed which is opposite of the relationship observed previously in horses[143] but consistent with the concept of higher latitudes experiencing greater seasonal extremes and thus having more pronounced photoperiod effects. Similar to previous reports, leptin values were highest in the month of October, at the end of the ACTH rise.[158] It is possible increased ACTH levels may have induced leptin resistance in an effort to prepare for winter. Fasted NEFA levels were higher during the winter months and may indicate decreased glucose utilization in favor of fat oxidation. The study findings suggest seasonality is an important source of metabolic trait variation; further evidence of this conclusion is the substantial reduction in farm level variance achieved for several traits with the inclusion of month as an explanatory variable. The horse, similar to other long day breeders, seems to switch to an insulin resistant state during the fall which would allow a horse to increase fat stores to be utilized throughout the winter when food sources may be limited and increased energy is needed for thermoregulation and to support development of a fetus in pregnant mares. Mild effects for diet and exercise parameters on metabolic trait variation were detected. Power may have been limited to detect an association with diet due to the one time sampling and limited power to detect associations with exercise due to only 15% of the population having received more than 3 hours of exercise per week.

6.1.5 Impact of EMS phenotype characterization findings and future directions

In summary, numerous factors beyond obesity and prior laminitis status were found to contribute to variation in metabolic traits, including seasonal effects, breed, age, sex, and diet. These findings obtained by examining equine metabolic trait variation in a large population of horses under varying environmental and individual conditions provide an explanation for the discrepancy in some conclusions drawn from previous studies regarding the EMS phenotype. However, these findings also draw attention to the difficulty in determining appropriate reference ranges for EMS diagnostic criteria in the presence of multiple sources of "normal" metabolic variation. A potential adjunct for improving EMS diagnosis prior to laminitis development and monitoring a horse's response to management would be to test multiple horses on the same property. A horse with values dissimilar from the rest of the herd may be at higher risk of laminitis development. Monitoring additional horses may be even more useful for monitoring progression of EMS. Instead of only monitoring the affected horse and questioning if the horse's values changed due to individual factors (progression/improvement of EMS) or due to environmental factors, one could monitor additional horses to determine if all of the horses experienced a similar change in test results due to farm level factors (season, diet).

Although parameters included in the analysis explained a moderate amount of variation in metabolic traits, the results also suggest additional unaccounted for sources of farm and individual horse level variance exist. It is likely that individual genetic differences exist which contributes to metabolic trait variation. The detection of breed differences in our study support this hypothesis. Additional environmental factors not measured in this study may also potentially impact metabolic trait variation. For example, chemicals present in the environment have endocrine disrupting capabilities and could be a potential unaccounted for source of variation in equine metabolic traits.[163]

A limitation of the present study is the one-time sampling of the population. The present study was not designed to ask questions such as: does metabolic trait variation and co-variation in horses with a prior history of laminitis vary with changes in body-weight? This should be a focus area of future research along with identifying additional sources of variation in metabolic traits in horses with a prior history of laminitis in both

obese and non-obese individuals. Further research is needed to identify additional environmental sources of metabolic trait variation and to dissect sources of individual level metabolic trait variation. Breed differences identified in Chapter 2 support the need for additional genetic studies, such as the GWAS performed in Chapter 5, and GWAS in multiple breeds to determine variants responsible for breed variation in metabolic traits.

The findings from Chapter 2 indicate both individual level factors (age, gender, genetics, obesity, and prior laminitis status) and environmental level factors (season, diet) are sources of equine metabolic trait variation. The metabolic trait profile of obese horses with and without a prior history of laminitis also vary providing further evidence adipose is more than just a fat storage site and the possibility an individual may have "healthy fat" vs "unhealthy fat".

6.2 Characterization of the equine response to an oral sugar challenge

Alteration of the equine incretin response (intestinal secretion of GLP-1 to stimulate insulin secretion in response to oral sugar) has been hypothesized to play a role in the pathogenesis of EMS however limited research has been performed in terms of testing this hypothesis. The objective of Chapter 3 was to characterize the equine response to an oral sugar challenge and identify factors associated with variation in this response. A longitudinal analysis was performed to test for association of different factors with variation in glucose, insulin, GLP-1 trajectories during an oral sugar test. The trajectory analysis is more informative than a traditional area under the curve summary statistic given the possibility for two different curves to have the same AUC resulting in a potential loss of important biological information. A number of factors were found to be associated with the incretin hormone GLP-1 and insulin/glucose dynamics during the equine response to an oral sugar challenge. In addition, factors associated with DPP4 activity, the major protease that breaks down GLP-1, were also identified.

6.2.1 Insulinemic response varies with breed, obesity, and prior laminitis status

Welsh ponies had a greater insulin response compared to Morgan horses suggesting a genetic basis for variation in the equine response to an oral sugar challenge. Obese horses with a prior history of laminitis also exhibited a greater insulin response. Triglycerides and leptin, biochemical factors found to be positively correlated with obesity and laminitis status in Chapter 2, also demonstrated a positive correlation with insulin response. Likewise, adiponectin demonstrated a negative correlation with the insulin response. The observed increased insulinemic response is likely due to an insulin resistant state in obese, prior laminitic horses, as the response is associated with biochemical variation consistent with insulin resistance.

6.2.2 Equine GLP-1 trajectories are not strongly correlated with the insulinemic response during an oral sugar test

GLP-1 trajectories were not strongly correlated with insulin trajectories. A potential explanation for the lack of correlation may be due to the use of a statistical model that assumes correlation of trajectories is the same in obese and non-obese horses and in horses with and without a prior history of laminitis. In humans the incretin effect accounts for 70% of the variation in the insulin response to an oral sugar challenge in normal individuals but only 30% percent in individuals affected with type 2 diabetes.[197] Estimation of trajectory correlation separately for obese and non-obese horses and in horses with and without a prior history of laminitis was not examined in Chapter 3 but should be investigated further in the future.

6.2.3 DPP4 activity is not associated with equine GLP-1 trajectories during an oral sugar test

An unexpected finding of Chapter 3 was the lack of association of DPP4 activity with equine GLP-1 trajectories given DPP4's role in degrading GLP-1 and the success of DPP4 inhibitors to modulate human GLP-1 levels and improve the incretin response. However, DPP4 activity was positively associated with the insulin response and negatively correlated with age and adiponectin levels, suggesting insulin resistance as the

mechanism for the increased insulin response and not modulation of GLP-1 levels by DPP4. The positive association of DPP4 activity with an insulin resistance type of response may reflect DPP4's role as an adipokine.[176]

6.2.4 Basal GLP-1 and GLP-1 secretory responses to oral glucose are reduced with features of EMS

Similar to decreased GLP-1 secretion observed with features of human metabolic syndrome (obesity and insulin resistance), basal GLP-1 and GLP-1 secretory responses to oral glucose are reduced with features of equine metabolic syndrome.[168,193–195] Haplotype variation in the promoter region of *GCG*, which encodes for the GLP-1 peptide, among other proteins was tested for association with GLP-1 and demonstrated a modest association. Haplotypes with higher basal GLP-1 levels tended to exhibit decreased GLP-1 responses to an oral sugar test. Testing additional horses would be necessary to confirm the haplotype as a causal source of GLP-1 variation. Although the role of variation in GLP-1 in the equine incretin response still remains to be questioned given the minimal correlation observed between insulin and GLP-1 trajectories. Future research should investigate a horse's response to an oral glucose stimulus compared with their response to an intravenous glucose stimulus to further assess the role of incretins in equine glucose intolerance.

6.3 Identification of candidate gene associated with metabolic syndrome phenotypes

6.3.1 Development of an improved linear mixed model to map polygenic traits in populations with familial relationships

The final objective of this thesis was to identify candidate genes associated with equine metabolic syndrome by performing a genome-wide association study of quantitative metabolic phenotypes. Human metabolic syndrome is a highly polygenic syndrome where numerous candidate genes have been identified but only explain a fraction of the heritability. Associated variants are typically of small effect size and require a large

sample size to detect associations. Horse populations do not randomly mate and experience substantial selection pressure, therefore it was hypothesized that a small number of moderate to large effect loci contribute to variation in metabolic traits and would be detectable with a relatively small population sample size (~300) in comparison to the very large (10,000+) sample sizes needed to detect candidate genes for human polygenic traits. The Morgan horse breed was selected for the GWAS given the reportedly higher prevalence of metabolic syndrome in this breed of horses. Estimation of "chip heritability" determined from a genomic relationship matrix calculated from all genotyped SNPs indicated metabolic traits were heritable, however initial genome-wide scans using standard linear mixed models failed to detect genome-wide significant SNPs, thus indicating causal variants were not well tagged by the SNPs or metabolic trait variation in the Morgan breed is due to several loci of small to moderate effect.

The standard LMM approach is based on single-locus tests combined with a diffuse, overall estimate of the genomic background based on all SNP markers. However, diffuse modeling of the polygenic term may not be appropriate for traits controlled by multiple loci of moderate effect. Recent LMM innovations aimed at improved modeling of the polygenic architecture by building a genomic relationship matrix (GRM) comprised of select SNPs associated with the phenotype have demonstrated increased power to map polygenic traits.[206–209] However, Widmer et al[208] cautioned building a GRM based on selected SNPs that well predict the phenotype performed poorly at controlling the type I error rate in the presence of familial relatedness and recommended inclusion of a second GRM comprised of all SNPs to achieve adequate type I error control. The drawback of this approach is increased computational time due to full-rank relatedness matrix inclusion. The Morgan horse population used for the GWAS performed in Chapter 5 featured familial relatedness, thus the objective of Chapter 4 was to develop and validate an improved linear mixed model for mapping polygenic traits in a population with familial relationships. The proposed model incorporated a Bayesian variable selection method to rank SNPs and a stepwise feature selection process to determine the optimal SNPs to model the random polygenic effect while including a random effect for each sampled herd or "familial cluster". The method was validated using the QTL-MAS 2010 dataset and real datasets for Morgan horse and Welsh pony height. The proposed method demonstrated increased power while controlling the false positive rate

and maintaining computational feasibility.

6.3.2 93 candidate loci identified for equine metabolic syndrome exhibit substantial overlap with human MetS candidate genes

The improved linear mixed model was applied to the 286 Morgan horse dataset genotyped on the Illumina SNP50 chip and imputed up to >800,000 SNPs to identify candidate genes for metabolic syndrome. 76 suggestive and 17 genome-wide significant candidate loci were identified and exhibited substantial overlap with human metabolic syndrome candidate genes such as *VEGFA*, *NRXN3*, *GRIK2*, and *TRIB2*. Additional interesting candidate genes include the top ranked loci for insulin levels following an oral sugar challenge, *ISL*, which encodes insulin enhancer protein that binds to the enhancer region of the insulin gene and may play an important role in regulating insulin gene expression. A insulin associated genetic variant near *AHR* which encodes the aryl hydrocarbon receptor, a ligand activated transcription factor known to bind endocrine disrupting chemicals such as polycyclic aromatic hydrocarbons and dioxins make it an interesting candidate gene to pursue given increased interest in the potential role of endocrine disrupting chemical in the pathophysiology of metabolic syndrome. Research already underway involves validation of the identified equine candidate genes in a second population and sequence analysis to identify the causal variants.

6.4 Summarizing statements

In summary, this work has increased knowledge and understanding of metabolic trait variation in horses and its relationship with the metabolic syndrome phenotype. A unifying theme of Chapter 2-5 was the repeated finding of similarities between horse and human metabolic syndrome and similar genetic architecture of complex traits in both species. Unexplained metabolic trait variation remains at the environmental level for horses after accounting for diet and exercise. Caloric intake and sedentary lifestyle have often been blamed for the obesity epidemic and obesity-related diseases in humans living in industrialized regions, however findings from this work indicate the possibility

of additional environmental sources of metabolic trait variation exist. Organic pollutants and endocrine disrupting chemicals have been proposed as playing a role in human obesity and type 2 diabetes and findings from this study warrant additional investigation. Identification of shared candidate genes for metabolic syndrome in both humans and horses both indicates the possibility of similar underlying pathophysiological mechanisms and provides opportunity for exploring similar underlying biological processes along with similar preventative and therapeutic management strategies.

References

[1] Kylin, E. Studies of the hypertension-hyperglycemia-hyperuricemia syndrome. *Zentralbl Inn Med*, 44:105–27, 1923.

[2] Vague, J. Sexual differentiation, a factor affecting the forms of obesity. *Presse Med*, 30:339–340, 1947.

[3] Avogaro, P., and Crepaldi, G. Essential hyperlipidemia, obesity and diabetes. *Diabetologia*, 1(137), 1965.

[4] Reaven, G. M. Role of insulin resistance in human disease. *Diabetes*, 37(12):1595–1607, 1988.

[5] Consultation, W. *Definition, diagnosis and classification of diabetes mellitus and its complications*, volume 1. Part, 1999.

[6] Balkau, B., and Charles, M. for the european group for the study of insulin resistance (egir) comment on the provisional report from the who consultation. *Diabetic Medicine*, 16:442–443, 1999.

[7] Einhorn, D., Reaven, G., Cobin, R., Ford, E., Ganda, O., Handelsman, Y., Hellman, R., Jellinger, P., Kendall, D., Krauss, R., et al. American college of endocrinology position statement on the insulin resistance syndrome. *Endocrine practice: official journal of the American College of Endocrinology and the American Association of Clinical Endocrinologists*, 9(3):237, 2003.

[8] Cleeman, J., Grundy, S., Becker, D., Clark, L., et al. Expert panel on detection, evaluation and treatment of high blood cholesterol in adults. executive summary of the third report of the national cholesterol education program (ncep) adult treatment panel (atp iii). *Jama*, 285(19):2486–2497, 2001.

[9] Alberti, K., Zimmet, P., and Shaw, J. Metabolic syndromea new world-wide definition. a consensus statement from the international diabetes federation. *Diabetic Medicine*, 23(5):469–480, 2006.

[10] Alberti, K., Eckel, R. H., Grundy, S. M., Zimmet, P. Z., Cleeman, J. I., Donato, K. A., Fruchart, J.-C., James, W. P. T., Loria, C. M., and Smith, S. C. Harmonizing the metabolic syndrome a joint interim statement of the international diabetes federation task force on epidemiology and prevention; national heart, lung, and blood institute; american heart association; world heart federation; international atherosclerosis society; and international association for the study of obesity. *Circulation*, 120(16):1640–1645, 2009.

[11] Desroches, S., and Lamarche, B. The evolving definitions and increasing prevalence of the metabolic syndrome. *Applied Physiology, Nutrition, and Metabolism*, 32(1):23–32, 2007.

[12] Kolovou, G. D., Anagnostopoulou, K. K., Salpea, K. D., and Mikhailidis, D. P. The prevalence of metabolic syndrome in various populations. *The American journal of the medical sciences*, 333(6):362–371, 2007.

[13] Zimmet, P. Z., and Alberti, G. The metabolic syndrome: Perhaps an etiologic mystery but far from a myth-where does the international diabetes federation stand. *Medscape Diabetes & Endocrinology*, 7(2), 2005.

[14] Hutley, L., and Prins, J. B. Fat as an endocrine organ: relationship to the metabolic syndrome. *The American journal of the medical sciences*, 330(6):280–289, 2005.

[15] Matsuzawa, Y., Funahashi, T., Kihara, S., and Shimomura, I. Adiponectin and metabolic syndrome. *Arteriosclerosis, thrombosis, and vascular biology*, 24(1):29–33, 2004.

[16] Kazumi, T., Kawaguchi, A., Sakai, K., Hirano, T., and Yoshino, G. Young men with high-normal blood pressure have lower serum adiponectin, smaller ldl size, and higher elevated heart rate than those with optimal blood pressure. *Diabetes care*, 25(6):971–976, 2002.

[17] Pischon, T., Girman, C. J., Hotamisligil, G. S., Rifai, N., Hu, F. B., and Rimm, E. B. Plasma adiponectin levels and risk of myocardial infarction in men. *Jama*, 291(14):1730–1737, 2004.

[18] Fumeron, F., Aubert, R., Siddiq, A., Betoulle, D., Péan, F., Hadjadj, S., Tichet, J., Wilpart, E., Chesnier, M.-C., Balkau, B., et al. Adiponectin gene polymorphisms and adiponectin levels are independently associated with the development of hyperglycemia during a 3-year period the epidemiologic data on the insulin resistance syndrome prospective study. *Diabetes*, 53(4):1150–1157, 2004.

[19] Lau, D. C., Dhillon, B., Yan, H., Szmitko, P. E., and Verma, S. Adipokines: molecular links between obesity and atheroslcerosis. *American Journal of Physiology-Heart and Circulatory Physiology*, 288(5):H2031–H2041, 2005.

[20] Carlyle, M., Jones, O. B., Kuo, J. J., and Hall, J. E. Chronic cardiovascular and renal actions of leptin role of adrenergic activity. *Hypertension*, 39(2):496–501, 2002.

[21] Considine, R. V., Sinha, M. K., Heiman, M. L., Kriauciunas, A., Stephens, T. W., Nyce, M. R., Ohannesian, J. P., Marco, C. C., McKee, L. J., Bauer, T. L., et al. Serum immunoreactive-leptin concentrations in normal-weight and obese humans. *New England Journal of Medicine*, 334(5):292–295, 1996.

[22] Banerjee, R. R., and Lazar, M. A. Resistin: molecular history and prognosis. *Journal of molecular medicine*, 81(4):218–226, 2003.

[23] Steppan, C. M., Bailey, S. T., Bhat, S., Brown, E. J., Banerjee, R. R., Wright, C. M., Patel, H. R., Ahima, R. S., and Lazar, M. A. The hormone resistin links obesity to diabetes. *Nature*, 409(6818):307–312, 2001.

[24] Cinti, S., Mitchell, G., Barbatelli, G., Murano, I., Ceresi, E., Faloia, E., Wang, S., Fortier, M., Greenberg, A. S., and Obin, M. S. Adipocyte death defines macrophage localization and function in adipose tissue of obese mice and humans. *Journal of lipid research*, 46(11):2347–2355, 2005.

[25] Weisberg, S. P., McCann, D., Desai, M., Rosenbaum, M., Leibel, R. L., Ferrante, A. W., et al. Obesity is associated with macrophage accumulation in adipose tissue. *Journal of Clinical Investigation*, 112(12):1796–1808, 2003.

[26] Xu, H., Uysal, K. T., Becherer, J. D., Arner, P., and Hotamisligil, G. S. Altered tumor necrosis factor-$\boldsymbol{\alpha}$ (tnf-$\boldsymbol{\alpha}$) processing in adipocytes and increased expression of transmembrane tnf-$\boldsymbol{\alpha}$ in obesity. *Diabetes*, 51(6):1876–1883, 2002.

[27] Hauner, H., Petruschke, T., Russ, M., Röhrig, K., and Eckel, J. Effects of tumour necrosis factor alpha (tnf$\boldsymbol{\alpha}$) on glucose transport and lipid metabolism of newly-differentiated human fat cells in cell culture. *Diabetologia*, 38(7):764–771, 1995.

[28] Zinman, B., Hanley, A. J., Harris, S. B., Kwan, J., and Fantus, I. G. Circulating tumor necrosis factor-$\boldsymbol{\alpha}$ concentrations in a native canadian population with high rates of type 2 diabetes mellitus 1. *The Journal of Clinical Endocrinology & Metabolism*, 84(1):272–278, 1999.

[29] Porter, M. H., Cutchins, A., Fine, J. B., Bai, Y., and Digirolamo, M. Effects of tnf-$\boldsymbol{\alpha}$ on glucose metabolism and lipolysis in adipose tissue and isolated fat-cell preparations. *Journal of Laboratory and Clinical Medicine*, 139(3):140–146, 2002.

[30] Banks, R., Forbes, M., Storr, M., Higginson, J., Thompson, D., Raynes, J., Illingworth, J., Perren, T., Selby, P., and Whicher, J. The acute phase protein response in patients receiving subcutaneous il-6. *Clinical & Experimental Immunology*, 102(1):217–223, 1995.

[31] Bulló, M., García-Lorda, P., Megias, I., and Salas-Salvadó, J. Systemic inflammation, adipose tissue tumor necrosis factor, and leptin expression. *Obesity Research*, 11(4):525–531, 2003.

[32] Kappes, A., and Löffler, G. Influences of ionomycin, dibutyryl-cycloamp and tumour necrosis factor-alpha on intracellular amount and secretion of apm1 in differentiating primary human preadipocytes. *Hormone and Metabolic Research*, 32(11/12):548–554, 2000.

[33] Kim, H.-J., Higashimori, T., Park, S.-Y., Choi, H., Dong, J., Kim, Y.-J., Noh, H.-L., Cho, Y.-R., Cline, G., Kim, Y.-B., et al. Differential effects of interleukin-6 and-10 on skeletal muscle and liver insulin action in vivo. *Diabetes*, 53(4):1060–1067, 2004.

[34] Rotter, V., Nagaev, I., and Smith, U. Interleukin-6 (il-6) induces insulin resistance in 3t3-l1 adipocytes and is, like il-8 and tumor necrosis factor-$\boldsymbol{\alpha}$, overexpressed in human fat cells from insulin-resistant subjects. *Journal of Biological Chemistry*, 278(46):45777–45784, 2003.

[35] Saleem, U., Khaleghi, M., Morgenthaler, N. G., Bergmann, A., Struck, J., Mosley Jr, T. H., and Kullo, I. J. Plasma carboxy-terminal provasopressin (copeptin): a novel marker of insulin resistance and metabolic syndrome. *The Journal of Clinical Endocrinology & Metabolism*, 94(7):2558–2564, 2009.

[36] Tsimikas, S., Willeit, J., Knoflach, M., Mayr, M., Egger, G., Notdurfter, M., Witztum, J. L., Wiedermann, C. J., Xu, Q., and Kiechl, S. Lipoprotein-associated phospholipase a2 activity, ferritin levels, metabolic syndrome, and 10-year cardiovascular and non-cardiovascular mortality: results from the bruneck study. *European heart journal*, 30(1):107–115, 2009.

[37] González, A. S., Guerrero, D. B., Soto, M. B., Díaz, S. P., Martinez-Olmos, M., and Vidal, O. Metabolic syndrome, insulin resistance and the inflammation markers c-reactive protein and ferritin. *European journal of clinical nutrition*, 60(6):802–809, 2006.

[38] Deepa, R., Velmurugan, K., Arvind, K., Sivaram, P., Sientay, C., Uday, S., and Mohan, V. Serum levels of interleukin 6, c-reactive protein, vascular cell adhesion molecule 1, and monocyte chemotactic protein 1 in relation to insulin resistance and glucose intolerancethe chennai urban rural epidemiology study (cures). *Metabolism*, 55(9):1232–1238, 2006.

[39] Guldiken, S., Demir, M., Arikan, E., Turgut, B., Azcan, S., Gerenli, M., and Tugrul, A. The levels of circulating markers of atherosclerosis and inflammation in subjects with different degrees of body mass index: soluble cd40 ligand and high-sensitivity c-reactive protein. *Thrombosis research*, 119(1):79–84, 2007.

[40] McLaughlin, T., Abbasi, F., Lamendola, C., Liang, L., Reaven, G., Schaaf, P., and Reaven, P. Differentiation between obesity and insulin resistance in the association with c-reactive protein. *Circulation*, 106(23):2908–2912, 2002.

[41] Wang, C. C. L., Goalstone, M. L., and Draznin, B. Molecular mechanisms of insulin resistance that impact cardiovascular biology. *Diabetes*, 53(11):2735–2740, 2004.

[42] Kahn, S. E., Hull, R. L., and Utzschneider, K. M. Mechanisms linking obesity to insulin resistance and type 2 diabetes. *Nature*, 444(7121):840–846, 2006.

[43] Flier, J. S., Underhill, L. H., and Eckel, R. H. Lipoprotein lipase. *New England Journal of Medicine*, 320(16):1060–1068, 1989.

[44] Eckel, R. H., Grundy, S. M., and Zimmet, P. Z. The metabolic syndrome. *The Lancet*, 365(9468):1415–1428, 2005.

[45] Roden, M., Price, T. B., Perseghin, G., Petersen, K. F., Rothman, D. L., Cline, G. W., and Shulman, G. I. Mechanism of free fatty acid-induced insulin resistance in humans. *Journal of Clinical Investigation*, 97(12):2859, 1996.

[46] Santomauro, A., Boden, G., Silva, M., Rocha, D. M., Santos, R. F., Ursich, M., Strassmann, P., and Wajchenberg, B. Overnight lowering of free fatty acids with acipimox improves insulin resistance and glucose tolerance in obese diabetic and nondiabetic subjects. *Diabetes*, 48(9):1836–1841, 1999.

[47] Boden, G., Lebed, B., Schatz, M., Homko, C., and Lemieux, S. Effects of acute changes of plasma free fatty acids on intramyocellular fat content and insulin resistance in healthy subjects. *Diabetes*, 50(7):1612–1617, 2001.

[48] Packard, C. Triacylglycerol-rich lipoproteins and the generation of small, dense low-density lipoprotein. *Biochemical Society Transactions*, 31(5):1066–1069, 2003.

[49] Kaur, J. A comprehensive review on metabolic syndrome. *Cardiology research and practice*, 2014, 2014.

[50] Lewis, G. Acute effects of insulin in the control of vldl production in humans. implications for the insulin-resistant state.[review][38 refs]. *Diabetes care*, 19(4):390–3, 1996.

[51] Ginsberg, H. N., Zhang, Y.-L., and Hernandez-Ono, A. Regulation of plasma triglycerides in insulin resistance and diabetes. *Archives of medical research*, 36(3):232–240, 2005.

[52] Manrique, C., Lastra, G., Gardner, M., and Sowers, J. R. The renin angiotensin aldosterone system in hypertension: roles of insulin resistance and oxidative stress. *Medical Clinics of North America*, 93(3):569–582, 2009.

[53] Schneider, D. J., and Sobel, B. E. Pai-1 and diabetes: a journey from the bench to the bedside. *Diabetes care*, 35(10):1961–1967, 2012.

[54] Chen, H., Montagnani, M., Funahashi, T., Shimomura, I., and Quon, M. J. Adiponectin stimulates production of nitric oxide in vascular endothelial cells. *Journal of Biological Chemistry*, 278(45):45021–45026, 2003.

[55] Khan, Q. A., Sola, S., and Khan, B. V. The metabolic syndrome: inflammation and endothelial dysfunction. *Hospital Physician*, 26:37, 2006.

[56] Baron, A. D. Vascular reactivity. *The American journal of cardiology*, 84(1):25–27, 1999.

[57] Perseghin, G., Ghosh, S., Gerow, K., and Shulman, G. I. Metabolic defects in lean nondiabetic offspring of niddm parents: a cross-sectional study. *Diabetes*, 46(6):1001–1009, 1997.

[58] Neel, J. V. Diabetes mellitus: a thrifty genotype rendered detrimental by progress? *American journal of human genetics*, 14(4):353, 1962.

[59] Tang, W., Hong, Y., Province, M. A., Rich, S. S., Hopkins, P. N., Arnett, D. K., Pankow, J. S., Miller, M. B., and Eckfeldt, J. H. Familial clustering for features of the metabolic syndrome the national heart, lung, and blood institute (nhlbi) family heart study. *Diabetes care*, 29(3):631–636, 2006.

[60] Povel, C., Boer, J., Reiling, E., and Feskens, E. Genetic variants and the metabolic syndrome: a systematic review. *Obesity Reviews*, 12(11):952–967, 2011.

[61] Frayling, T. M., Timpson, N. J., Weedon, M. N., Zeggini, E., Freathy, R. M., Lindgren, C. M., Perry, J. R., Elliott, K. S., Lango, H., Rayner, N. W., et al. A common variant in the fto gene is associated with body mass index and predisposes to childhood and adult obesity. *Science*, 316(5826):889–894, 2007.

[62] Sladek, R., Rocheleau, G., Rung, J., Dina, C., Shen, L., Serre, D., Boutin, P., Vincent, D., Belisle, A., Hadjadj, S., et al. A genome-wide association study identifies novel risk loci for type 2 diabetes. *Nature*, 445(7130):881–885, 2007.

[63] Huth, C., Illig, T., Herder, C., Gieger, C., Grallert, H., Vollmert, C., Rathmann, W., Hamid, Y. H., Pedersen, O., Hansen, T., et al. Joint analysis of individual participants' data from 17 studies on the association of the il6 variant-174g¿ c with circulating glucose levels, interleukin-6 levels, and body mass index. *Annals of medicine*, 41(2):128–138, 2009.

[64] Grallert, H., Sedlmeier, E.-M., Huth, C., Kolz, M., Heid, I. M., Meisinger, C., Herder, C., Strassburger, K., Gehringer, A., Haak, M., et al. Apoa5 variants and metabolic syndrome in caucasians. *Journal of lipid research*, 48(12):2614–2621, 2007.

[65] Hegele, R. A., Connelly, P. W., Hanley, A. J., Sun, F., Harris, S. B., and Zinman, B. Common genomic variation in the apoc3 promoter associated with variation in plasma lipoproteins. *Arteriosclerosis, thrombosis, and vascular biology*, 17(11):2753–2758, 1997.

[66] Thompson, A., Di Angelantonio, E., Sarwar, N., Erqou, S., Saleheen, D., Dullaart, R. P., Keavney, B., Ye, Z., and Danesh, J. Association of cholesteryl ester transfer protein genotypes with cetp mass and activity, lipid levels, and coronary risk. *Jama*, 299(23):2777–2788, 2008.

[67] Speakman, J., Hambly, C., Mitchell, S., and Król, E. The contribution of animal models to the study of obesity. *Laboratory animals*, 42(4):413–432, 2008.

[68] Cottrell, E. C., and Ozanne, S. E. Developmental programming of energy balance and the metabolic syndrome. *Proceedings of the Nutrition Society*, 66(02):198–206, 2007.

[69] Oron-Herman, M., Kamari, Y., Grossman, E., Yeger, G., Peleg, E., Shabtay, Z., Shamiss, A., and Sharabi, Y. Metabolic syndrome: comparison of the two commonly used animal models. *American journal of hypertension*, 21(9):1018–1022, 2008.

[70] Gajda, A. M., Pellizzon, M. A., Ricci, M. R., and Ulman, E. A. Diet-induced metabolic syndrome in rodent models. *Animal Lab News*, 74:775–793, 2007.

[71] Sone, H., Takahashi, A., Iida, K., and Yamada, N. Disease model: hyperinsulinemia and insulin resistance part b–polygenic and other animal models. *Trends in molecular medicine*, 7(8):373–376, 2001.

[72] Bergman, R. N., Kim, S. P., Catalano, K. J., Hsu, I. R., Chiu, J. D., Kabir, M., Hucking, K., and Ader, M. Why visceral fat is bad: mechanisms of the metabolic syndrome. *Obesity*, 14(S2):16S–19S, 2006.

[73] Russell, J. C., and Proctor, S. D. Small animal models of cardiovascular disease: tools for the study of the roles of metabolic syndrome, dyslipidemia, and atherosclerosis. *Cardiovascular pathology*, 15(6):318–330, 2006.

[74] Xi, S., Yin, W., Wang, Z., Kusunoki, M., Lian, X., Koike, T., Fan, J., and Zhang, Q. A minipig model of high-fat/high-sucrose diet-induced diabetes and atherosclerosis. *International journal of experimental pathology*, 85(4):223–231, 2004.

[75] Johnson, P. J. The equine metabolic syndrome: peripheral cushing's syndrome. *Veterinary Clinics of North America: Equine Practice*, 18(2):271–293, 2002.

[76] Frank, N., Geor, R., Bailey, S., Durham, A., and Johnson, P. Equine metabolic syndrome. *Journal of Veterinary Internal Medicine*, 24(3):467–475, 2010.

[77] Treiber, K. H., Kronfeld, D. S., Hess, T. M., Byrd, B. M., Splan, R. K., and Staniar, W. B. Evaluation of genetic and metabolic predispositions and nutritional risk factors for pasture-associated laminitis in ponies. *Journal of the American Veterinary Medical Association*, 228(10):1538–1545, 2006.

[78] Carter, R. A., Treiber, K., Geor, R., Douglass, L., and Harris, P. A. Prediction of incipient pasture-associated laminitis from hyperinsulinaemia, hyperleptinaemia and generalised and localised obesity in a cohort of ponies. *Equine veterinary journal*, 41(2):171–178, 2009.

[79] Frank, N., Elliott, S. B., Brandt, L. E., and Keisler, D. H. Physical characteristics, blood hormone concentrations, and plasma lipid concentrations in obese horses with insulin resistance. *Journal of the American Veterinary Medical Association*, 228(9):1383–1390, 2006.

[80] Cartmill, J., Thompson, D., Storer, W., Gentry, L., and Huff, N. Endocrine responses in mares and geldings with high body condition scores grouped by high vs. low resting leptin concentrations. *Journal of animal science*, 81(9):2311–2321, 2003.

[81] Bailey, S. R., Habershon-Butcher, J. L., Ransom, K. J., Elliott, J., and Menzies-Gow, N. J. Hypertension and insulin resistance in a mixed-breed population of ponies predisposed to laminitis. *American journal of veterinary research*, 69(1):122–129, 2008.

[82] Gentry, L., Thompson, D., Gentry, G., Davis, K., Godke, R., and Cartmill, J. The relationship between body condition, leptin, and reproductive and hormonal characteristics of mares during the seasonal anovulatory period. *Journal of animal science*, 80(10):2695–2703, 2002.

[83] Vick, M., Sessions, D., Murphy, B., Kennedy, E., Reedy, S., and Fitzgerald, B. Obesity is associated with altered metabolic and reproductive activity in the mare: effects of metformin on insulin sensitivity and reproductive cyclicity. *Reproduction, Fertility and Development*, 18(6):609–617, 2006.

[84] Vick, M., Adams, A., Murphy, B., Sessions, D., Horohov, D., Cook, R., Shelton, B., and Fitzgerald, B. Relationships among inflammatory cytokines, obesity, and insulin sensitivity in the horse. *Journal of animal science*, 85(5):1144–1155, 2007.

[85] McCue, M. E., Geor, R. J., and Schultz, N. Equine metabolic syndrome: A complex disease influenced by genetics and the environment. *Journal of Equine Veterinary Science*, 35(5):367–375, 2015.

[86] Treiber, K., Hess, T., Kronfeld, D., Boston, R., Geor, R., and Harris, P. Insulin resistance and compensation in laminitis-predisposed ponies characterized by the minimal model. *Pferdeheilkunde*, 21:91–92, 2005.

[87] Bamford, N., Potter, S., Harris, P., and Bailey, S. Breed differences in insulin sensitivity and insulinemic responses to oral glucose in horses and ponies of moderate body condition score. *Domestic animal endocrinology*, 47:101–107, 2014.

[88] Treiber, K. H., Kronfeld, D. S., Hess, T. M., Boston, R. C., and Harris, P. A. Use of proxies and reference quintiles obtained from minimal model analysis for determination of insulin sensitivity and pancreatic beta-cell responsiveness in horses. *American journal of veterinary research*, 66(12):2114–2121, 2005.

[89] Bailey, S. R., Menzies-Gow, N. J., Harris, P. A., Habershon-Butcher, J. L., Crawford, C., Berhane, Y., Boston, R. C., and Elliott, J. Effect of dietary fructans and dexamethasone administration on the insulin response of ponies predisposed to laminitis. *Journal of the American Veterinary Medical Association*, 231(9):1365–1373, Nov 2007.

[90] Toft-Nielsen, M.-B., Damholt, M. B., Madsbad, S., Hilsted, L. M., Hughes, T. E., Michelsen, B. K., and Holst, J. J. Determinants of the impaired secretion of glucagon-like peptide-1 in type 2 diabetic patients. *The Journal of Clinical Endocrinology & Metabolism*, 86(8):3717–3723, 2001.

[91] Shanik, M. H., Xu, Y., Škrha, J., Dankner, R., Zick, Y., and Roth, J. Insulin resistance and hyperinsulinemia is hyperinsulinemia the cart or the horse? *Diabetes care*, 31(Supplement 2):S262–S268, 2008.

[92] de Graaf-Roelfsema, E. Glucose homeostasis and the enteroinsular axis in the horse: A possible role in equine metabolic syndrome. *The Veterinary Journal*, 199(1):11–18, 2014.

[93] Frank, N., and Tadros, E. Insulin dysregulation. *Equine veterinary journal*, 46(1):103–112, 2014.

[94] Becvarova, I., and Pleasant, R. S. Managing obesity in pasture-based horses. *Compend. Contin. Educ. Vet*, 34, 2012.

[95] Morgan, R., McGowan, T., and McGowan, C. Prevalence and risk factors for hyperinsulinaemia in ponies in queensland, australia. *Australian veterinary journal*, 92(4):101–106, 2014.

[96] Suagee, J., Corl, B., Crisman, M., Pleasant, R., Thatcher, C., and Geor, R. Relationships between body condition score and plasma inflammatory cytokines, insulin, and lipids in a mixed population of light-breed horses. *Journal of Veterinary Internal Medicine*, 27(1):157–163, 2013.

[97] Treiber, K., Carter, R., Gay, L., Williams, C., and Geor, R. Inflammatory and redox status of ponies with a history of pasture-associated laminitis. *Veterinary immunology and immunopathology*, 129(3):216–220, 2009.

[98] Carter, R. A., McCutcheon, L. J., George, L. A., Smith, T. L., Frank, N., and Geor, R. J. Effects of diet-induced weight gain on insulin sensitivity and plasma hormone and lipid concentrations in horses. *American journal of veterinary research*, 70(10):1250–1258, 2009.

[99] Van Weyenberg, S., Hesta, M., Buyse, J., and Janssens, G. The effect of weight loss by energy restriction on metabolic profile and glucose tolerance in ponies. *Journal of animal physiology and animal nutrition*, 92(5):538–545, 2008.

[100] Quinn, R. W., Burk, A. O., Hartsock, T. G., Petersen, E. D., Whitley, N. C., Treiber, K. H., and Boston, R. C. Insulin sensitivity in thoroughbred geldings: effect of weight gain, diet, and exercise on insulin sensitivity in thoroughbred geldings. *Journal of equine veterinary science*, 28(12):728–738, 2008.

[101] Wooldridge, A. A., Edwards, H. G., Plaisance, E. P., Applegate, R., Taylor, D. R., Taintor, J., Zhong, Q., and Judd, R. L. Evaluation of high–molecular weight adiponectin in horses. *American journal of veterinary research*, 73(8):1230–1240, 2012.

[102] Kearns, C. F., McKeever, K. H., Roegner, V., Brady, S. M., and Malinowski, K. Adiponectin and leptin are related to fat mass in horses. *The Veterinary Journal*, 172(3):460–465, 2006.

[103] Wray, H., Elliott, J., Bailey, S., Harris, P., and Menzies-Gow, N. Plasma concentrations of inflammatory markers in previously laminitic ponies. *Equine veterinary journal*, 45(5):546–551, 2013.

[104] Ungru, J., Blüher, M., Coenen, M., Raila, J., Boston, R., and Vervuert, I. Effects of body weight reduction on blood adipokines and subcutaneous adipose tissue adipokine mrna expression profiles in obese ponies. *Veterinary Record*, pages vetrec–2012, 2012.

[105] Johnson, P. J., Wiedmeyer, C. E., LaCarrubba, A., Ganjam, V., Messer, I., and Nat, T. Laminitis and the equine metabolic syndrome. *Veterinary Clinics of North America: Equine Practice*, 26(2):239–255, 2010.

[106] Asplin, K. E., Sillence, M. N., Pollitt, C. C., and McGowan, C. M. Induction of laminitis by prolonged hyperinsulinaemia in clinically normal ponies. *The veterinary journal*, 174(3):530–535, 2007.

[107] de Laat, M. A., McGowan, C., Sillence, M., and Pollitt, C. Equine laminitis: induced by 48 h hyperinsulinaemia in standardbred horses. *Equine veterinary journal*, 42(2):129–135, 2010.

[108] Luke, D. A. Multilevel modeling: Quantitative applications in the social sciences. *London: A sage university paper series*, 2004.

[109] Raudenbush, S. W., and Bryk, A. S. *Hierarchical linear models: Applications and data analysis methods*, volume 1. Sage, 2002.

[110] Longford, N. T. *Random coefficient models*. Springer, 1995.

[111] Pinheiro, J. C., and Bates, D. M. *Mixed-effects models in S and S-PLUS*. Springer, 2000.

[112] McArdle, J. J., and Epstein, D. Latent growth curves within developmental structural equation models. *Child development*, pages 110–133, 1987.

[113] Snijders, T., and Bosker, R. *Multilevel analysis: an introduction to basic and advanced multilevel modelingSage Publications*. 1999.

[114] Akaike, H. Factor analysis and aic. *Psychometrika*, 52(3):317–332, 1987.

[115] Treiber, K., Kronfeld, D., Hess, T., Byrd, B., and Splan, R. Pre-laminitic metabolic syndrome in genetically predisposed ponies involves compensated insulin resistance. *Journal of Animal Physiology and Animal Nutrition*, 89(11-12):430–431, 2005.

[116] Treiber, K., Hess, T., Kronfeld, D., Boston, R., Geor, R., Friere, M., Silva, A., and Harris, P. Glucose dynamics during exercise: dietary energy sources affect minimal model parameters in trained arabian geldings during endurance exercise. *Equine Veterinary Journal*, 38(S36):631–636, 2006.

[117] Fowler, F. *Survey Research Methods (Applied Social Research Methods)*. Sage Publications, Inc, 2001.

[118] Henneke, D., Potter, G., Kreider, J., and Yeates, B. Relationship between condition score, physical measurements and body fat percentage in mares. *Equine veterinary journal*, 15(4):371–372, 1983.

[119] Martinson, K., Coleman, R., Rendahl, A., Fang, Z., and McCue, M. Estimation of body weight and development of a body weight score for adult equids using morphometric measurements. *Journal of animal science*, 92(5):2230–2238, 2014.

[120] Carroll, C., and Huntington, P. Body condition scoring and weight estimation of horses. *Equine veterinary journal*, 20(1):41–45, 1988.

[121] Borer-Weir, K. E., Bailey, S. R., Menzies-Gow, N. J., Harris, P. A., and Elliott, J. Evaluation of a commercially available radioimmunoassay and species-specific elisas for measurement of high concentrations of insulin in equine serum. *American journal of veterinary research*, 73(10):1596–1602, 2012.

[122] Panzani, S., Comin, A., Galeati, G., Romano, G., Villani, M., Faustini, M., and Veronesi, M. How type of parturition and health status influence hormonal and metabolic profiles in newborn foals. *Theriogenology*, 77(6):1167–1177, 2012.

[123] Perkins, G., Lamb, S., Erb, H., Schanbacher, B., Nydam, D., and Divers, T. Plasma adrenocorticotropin (acth) concentrations and clinical response in horses treated for equine cushing's disease with cyproheptadine or pergolide. *Equine veterinary journal*, 34(7):679–685, 2002.

[124] Fitzgerald, B. P., and McManus, C. J. Photoperiodic versus metabolic signals as determinants of seasonal anestrus in the mare. *Biology of reproduction*, 63(1):335–340, 2000.

[125] McManus, C., and Fitzgerald, B. Effect of daily clenbuterol and exogenous melatonin treatment on body fat, serum leptin and the expression of seasonal anestrus in the mare. *Animal reproduction science*, 76(3):217–230, 2003.

[126] Glunk, E., and Siciliano, P. Effect of restricted grazing on dry matter intake rate. *Journal of Equine Veterinary Science*, 31(5):296–297, 2011.

[127] Longland, A., Barfoot, C., and Harris, P. The effect of wearing a grazing muzzle vs. not wearing a grazing muzzle on intakes of spring, summer and autumn pastures by ponies. In *Forages and grazing in horse nutrition*, pages 185–186. Springer, 2012.

[128] Goldstein, H. *Multilevel Statistical Models*. A Hodder Arnold Publication, 2002.

[129] Butler, D., Cullis, B. R., Gilmour, A., and Gogel, B. Asreml-r reference manual. *Brisbane: Queensland Department of Primary Industries and Fisheries*, 2007.

[130] Lipovetsky, S., and Conklin, M. Analysis of regression in game theory approach. *Applied Stochastic Models in Business and Industry*, 17(4):319–330, 2001.

[131] Cohen, J. *Statistical power analysis for the behavioral sciences*. Academic press, 2013.

[132] Park, K.-G., Park, K. S., Kim, M.-J., Kim, H.-S., Suh, Y.-S., Ahn, J. D., Park, K.-K., Chang, Y.-C., and Lee, I.-K. Relationship between serum adiponectin and leptin concentrations and body fat distribution. *Diabetes research and clinical practice*, 63(2):135–142, 2004.

[133] Burns, T., Geor, R., Mudge, M., McCutcheon, L., Hinchcliff, K., and Belknap, J. Proinflammatory cytokine and chemokine gene expression profiles in subcutaneous and visceral adipose tissue depots of insulin-resistant and insulin-sensitive light breed horses. *Journal of Veterinary Internal Medicine*, 24(4):932–939, 2010.

[134] Thamer, C., Machann, J. r., Staiger, H., Mu ssig, K., Schwenzer, N., Ludescher, B., Machicao, F., Claussen, C., Fritsche, A., Schick, F., et al. Interscapular fat is strongly associated with insulin resistance. *The Journal of Clinical Endocrinology & Metabolism*, 95(10):4736–4742, 2010.

[135] Hasan-Ali, H., El-Mottaleb, N. A. A., Hamed, H. B., and Abd-Elsayed, A. Serum adiponectin and leptin as predictors of the presence and degree of coronary atherosclerosis. *Coronary artery disease*, 22(4):264–269, 2011.

[136] Petersen, J. L., Mickelson, J. R., Rendahl, A. K., Valberg, S. J., Andersson, L. S., Axelsson, J., Bailey, E., Bannasch, D., Binns, M. M., Borges, A. S., et al. Genome-wide analysis reveals selection for important traits in domestic horse breeds. 2013.

[137] Dilger, A. C., Spurlock, M. E., Grant, A. L., and Gerrard, D. E. Myostatin null mice respond differently to dietary-induced and genetic obesity. *Animal Science Journal*, 81(5):586–593, 2010.

[138] Guo, T., Jou, W., Chanturiya, T., Portas, J., Gavrilova, O., McPherron, A. C., et al. Myostatin inhibition in muscle, but not adipose tissue, decreases fat mass and improves insulin sensitivity. *PloS one*, 4(3):e4937, 2009.

[139] Hamrick, M., Pennington, C., Webb, C., and Isales, C. Resistance to body fat gain in 'double-muscled'mice fed a high-fat diet. *International journal of obesity*, 30(5):868–870, 2006.

[140] Lin, J., Arnold, H. B., Della-Fera, M. A., Azain, M. J., Hartzell, D. L., and Baile, C. A. Myostatin knockout in mice increases myogenesis and decreases adipogenesis. *Biochemical and biophysical research communications*, 291(3):701–706, 2002.

[141] McPherron, A. C., Lee, S.-J., et al. Suppression of body fat accumulation in myostatin-deficient mice. *The Journal of clinical investigation*, 109(109 (5)):595–601, 2002.

[142] Zhang, C., McFarlane, C., Lokireddy, S., Masuda, S., Ge, X., Gluckman, P. D., Sharma, M., and Kambadur, R. Inhibition of myostatin protects against diet-induced obesity by enhancing fatty acid oxidation and promoting a brown adipose phenotype in mice. *Diabetologia*, 55(1):183–193, 2012.

[143] McFarlane, D., Paradis, M., Zimmel, D., Sykes, B., Brorsen, B., Sanchez, A., and Vainio, K. The effect of geographic location, breed, and pituitary dysfunction on seasonal adrenocorticotropin and $\boldsymbol{\alpha}$-melanocyte-stimulating hormone plasma concentrations in horses. *Journal of Veterinary Internal Medicine*, 25(4):872–881, 2011.

[144] Donaldson, M. T., McDonnell, S. M., Schanbacher, B. J., Lamb, S. V., McFarlane, D., and Beech, J. Variation in plasma adrenocorticotropic hormone concentration and dexamethasone suppression test results with season, age, and sex in healthy ponies and horses. *Journal of veterinary internal medicine*, 19(2):217–222, 2005.

[145] Shimizu, H., Shimomura, Y., Hayashi, R., Ohtani, K., Sato, N., Futawatari, T., and Mori, M. Serum leptin concentration is associated with total body fat mass, but not abdominal fat distribution. *International journal of obesity and related metabolic disorders: journal of the International Association for the Study of Obesity*, 21(7):536–541, 1997.

[146] Seidell, J. C., Cigolini, M., Charzewska, J., Ellsinger, B.-M., Björntorp, P., Hautvast, J. G., and Szostak, W. Fat distribution and gender differences in serum lipids in men and women from four european communities. *Atherosclerosis*, 87(2):203–210, 1991.

[147] Marshall, J. A., Grunwald, G. K., Donahoo, W. T., Scarbro, S., and Shetterly, S. M. Percent body fat and lean mass explain the gender difference in leptin: Analysis and interpretation of leptin in hispanic and non-hispanic white adults. *Obesity Research*, 8(8):543–552, 2000.

[148] Nielsen, B. D., O'Connor-Robison, C. I., Spooner, H. S., and Shelton, J. Glycemic and insulinemic responses are affected by age of horse and method of feed processing. *Journal of Equine Veterinary Science*, 30(5):249–258, 2010.

[149] Place, N., McGowan, C., Lamb, S., Schanbacher, B., McGowan, T., and Walsh, D. Seasonal variation in serum concentrations of selected metabolic hormones in horses. *Journal of Veterinary Internal Medicine*, 24(3):650–654, 2010.

[150] Frank, N., Elliott, S. B., Chameroy, K., Toth, F., Chumbler, N., and McClamroch, R. Association of season and pasture grazing with blood hormone and metabolite concentrations in horses with presumed pituitary pars intermedia dysfunction. *Journal of Veterinary Internal Medicine*, 24(5):1167–1175, 2010.

[151] Schreiber, C. M., Stewart, A. J., Kwessi, E., Behrend, E. N., Wright, J. C., Kemppainen, R. J., and Busch, K. A. Seasonal variation in results of diagnostic tests for pituitary pars intermedia dysfunction in older, clinically normal geldings. *Journal of the American Veterinary Medical Association*, 241(2):241–248, 2012.

[152] Michel, C., Chastel, O., and Bonnet, X. Ambient temperature and pregnancy influence cortisol levels in female guinea pigs and entail long-term effects on the stress response of their offspring. *General and comparative endocrinology*, 171(3):275–282, 2011.

[153] Marchal, J., Dorieux, O., Haro, L., Aujard, F., and Perret, M. Characterization of blood biochemical markers during aging in the grey mouse lemur (microcebus murinus): impact of gender and season. *BMC veterinary research*, 8(1):211, 2012.

[154] Behall, K., Scholfield, D., Hallfrisch, J., Kelsay, J., and Reiser, S. Seasonal variation in plasma glucose and hormone levels in adult men and women. *The American journal of clinical nutrition*, 40(6):1352–1356, 1984.

[155] Borer, K., Bailey, S., MENZIES-GOW, N., Harris, P., and Elliott, J. Use of proxy measurements of insulin sensitivity and insulin secretory response to distinguish between normal and previously laminitic ponies. *Equine Veterinary Journal*, 44(4):444–448, 2012.

[156] Thornton, P., Wright, P. A., Sacra, P., and Goodier, T. The ferret, mustela putorius furo, as a new species in toxicology. *Laboratory animals*, 13(2):119–124, 1979.

[157] Nieminen, P., Asikainen, J., and Hyvärinen, H. Effects of seasonality and fasting on the plasma leptin and thyroxin levels of the raccoon dog (nyctereutes procyonoides) and the blue fox (alopex lagopus). *Journal of Experimental Zoology*, 289(2):109–118, 2001.

[158] Elliott, S. Effects of pituitary pars intermedia dysfunction (ppid), season, and pasture diet on blood adrenocorticotropic hormone and metabolite concentrations in horses. *Masters Theses*, 2010.

[159] Cahill, S., Tuplin, E., and Holahan, M. R. Circannual changes in stress and feeding hormones and their effect on food-seeking behaviors. *Frontiers in neuroscience*, 7, 2013.

[160] George, L. A., Staniar, W. B., Cubitt, T. A., Treiber, K. H., Harris, P. A., and Geor, R. J. Evaluation of the effects of pregnancy on insulin sensitivity, insulin secretion, and glucose dynamics in thoroughbred mares. *American journal of veterinary research*, 72(5):666–674, 2011.

[161] Hoffman, R., Kronfeld, D., Cooper, W., and Harris, P. Glucose clearance in grazing mares is affected by diet, pregnancy, and lactation. *Journal of animal science*, 81(7):1764–1771, 2003.

[162] Kaneko, T., Wang, P.-Y., Tawata, M., and Sato, A. Low carbohydrate intake before oral glucose-tolerance tests. *The Lancet*, 352(9124):289, 1998.

[163] Lubrano, C., Genovesi, G., Specchia, P., Costantini, D., Mariani, S., Petrangeli, E., Lenzi, A., and Gnessi, L. Obesity and metabolic comorbidities: environmental diseases? *Oxidative medicine and cellular longevity*, 2013:640673, Mar 2013.

[164] Baggio, L. L., and Drucker, D. J. Biology of incretins: Glp-1 and gip. *Gastroenterology*, 132(6):2131–2157, 2007.

[165] Kieffer, T. J., McIntosh, C., and Pederson, R. A. Degradation of glucose-dependent insulinotropic polypeptide and truncated glucagon-like peptide 1 in vitro and in vivo by dipeptidyl peptidase iv. *Endocrinology*, 136(8):3585–3596, 1995.

[166] Muscelli, E., Mari, A., Casolaro, A., Camastra, S., Seghieri, G., Gastaldelli, A., Holst, J. J., and Ferrannini, E. Separate impact of obesity and glucose tolerance on the incretin effect in normal subjects and type 2 diabetic patients. *Diabetes*, 57(5):1340–1348, 2008.

[167] Tura, A., Muscelli, E., Gastaldelli, A., Ferrannini, E., and Mari, A. Altered pattern of the incretin effect as assessed by modelling in individuals with glucose tolerance ranging from normal to diabetic. *Diabetologia*, 57(6):1199–1203, 2014.

[168] Zhang, F., Tang, X., Cao, H., Lü, Q., Li, N., Liu, Y., Zhang, X., Zhang, Y., Cao, M., Wan, J., et al. Impaired secretion of total glucagon-like peptide-1 in people with impaired fasting glucose combined impaired glucose tolerance. *International journal of medical sciences*, 9(7):574, 2012.

[169] J Neumiller, J. Incretin pharmacology: a review of the incretin effect and current incretin-based therapies. *Cardiovascular & Hematological Agents in Medicinal Chemistry (Formerly Current Medicinal Chemistry-Cardiovascular & Hematological Agents)*, 10(4):276–288, 2012.

[170] Petersen, J., Mickelson, J., and McCue, M. The identification of signatures of selection in modern horse breeds using genome-wide snp data. In *Plant & Animal Genome XIX Conference, San Diego, CA*, 2011.

[171] Nian, M., Drucker, D. J., and Irwin, D. Divergent regulation of human and rat proglucagon gene promoters in vivo. *American Journal of Physiology-Gastrointestinal and Liver Physiology*, 277(4):G829–G837, 1999.

[172] Smit, A., Hubley, R., and Green, P. 1996–2010. repeatmasker open-3.0. *URL: http://www. repeatmasker. org*.

[173] Allison, D. B., Paultre, F., Maggio, C., Mezzitis, N., and Pi-Sunyer, F. X. The use of areas under curves in diabetes research. *Diabetes care*, 18(2):245–250, 1995.

[174] Frøslie, K. F., Røislien, J., Qvigstad, E., Godang, K., Bollerslev, J., Voldner, N., Henriksen, T., and Veierød, M. B. Shape information from glucose curves: Functional data analysis compared with traditional summary measures. *BMC medical research methodology*, 13(1):6, 2013.

[175] Kirino, Y., Sei, M., Kawazoe, K., Minakuchi, K., and Sato, Y. Plasma dipeptidyl peptidase 4 activity correlates with body mass index and the plasma adiponectin concentration in healthy young people. *Endocrine journal*, 59(10):949–953, 2011.

[176] Lamers, D., Famulla, S., Wronkowitz, N., Hartwig, S., Lehr, S., Ouwens, D. M., Eckardt, K., Kaufman, J. M., Ryden, M., Müller, S., et al. Dipeptidyl peptidase 4 is a novel adipokine potentially linking obesity to the metabolic syndrome. *Diabetes*, 60(7):1917–1925, 2011.

[177] Korosi, J., McIntosh, C. H., Pederson, R. A., Demuth, H.-U., Habener, J. F., Gingerich, R., Egan, J. M., Elahi, D., and Meneilly, G. S. Effect of aging and diabetes on the enteroinsular axis. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*, 56(9):M575–M579, 2001.

[178] Stankovic, M., Vlahovic, P., Avramovic, V., and Todorovic, M. Distribution of dipeptidyl peptidase iv in patients with chronic tonsillitis. *Clinical and Vaccine Immunology*, 15(5):794–798, 2008.

[179] Drozdowski, L., and Thomson, A. B. Aging and the intestine. *World journal of gastroenterology*, 12(47):7578, 2006.

[180] Ionescu, E., Sauter, J., and Jeanrenaud, B. Abnormal oral glucose tolerance in genetically obese (fa/fa) rats. *Am J Physiol*, 248(5 Pt 1):E500–6, 1985.

[181] Scheen, A., Paquot, N., Letiexhe, M., Paolisso, G., Castillo, M., and Lefebvre, P. Glucose metabolism in obese subjects: lessons from ogtt, ivgtt and clamp studies. *International journal of obesity and related metabolic disorders: journal of the International Association for the Study of Obesity*, 19:S14–20, 1995.

[182] Jenkins, D. J., and Jenkins, A. L. Dietary fiber and the glycemic response. *Experimental Biology and Medicine*, 180(3):422–431, 1985.

[183] Borer, K., Bailey, S., Menzies-Gow, N., Harris, P., and Elliott, J. Effect of feeding glucose, fructose, and inulin on blood glucose and insulin concentrations in normal ponies and those predisposed to laminitis. *Journal of animal science*, 90(9):3003–3011, 2012.

[184] Dresner, A., Laurent, D., Marcucci, M., Griffin, M. E., Dufour, S., Cline, G. W., Slezak, L. A., Andersen, D. K., Hundal, R. S., Rothman, D. L., et al. Effects of free fatty acids on glucose transport and irs-1–associated phosphatidylinositol 3-kinase activity. *The Journal of clinical investigation*, 103(2):253–259, 1999.

[185] Lam, T. K., Van de Werve, G., and Giacca, A. Free fatty acids increase basal hepatic glucose production and induce hepatic insulin resistance at different sites. *American Journal of Physiology-Endocrinology And Metabolism*, 284(2):E281–E290, 2003.

[186] Staehr, P., Hother-Nielsen, O., Landau, B. R., Chandramouli, V., Holst, J. J., and Beck-Nielsen, H. Effects of free fatty acids per se on glucose production, gluconeogenesis, and glycogenolysis. *Diabetes*, 52(2):260–267, 2003.

[187] Chiles, R., Tzagournis, M., and Catalano, E. Excessive serum insulin response to oral glucose in obesity and mild diabetes study of 501 patients. *Diabetes*, 19(6):458–464, 1970.

[188] Farrant, P., Neville, R., and Stewart, G. Insulin release in response to oral glucose in obesity: The effect of reduction of body weight. *Diabetologia*, 5(3):198–200, 1969.

[189] Laws, A., and Reaven, G. Evidence for an independent relationship between insulin resistance and fasting plasma hdl-cholesterol, triglyceride and insulin concentrations. *Journal of internal medicine*, 231(1):25–30, 1992.

[190] Lee, C.-Y., Lee, C.-H., Tsai, S., Huang, C.-T., Wu, M.-T., Tai, S.-Y., Lin, F.-F., Chao, N.-C., and Chang, C.-J. Association between serum leptin and adiponectin levels with risk of insulin resistance and impaired glucose tolerance in non-diabetic women. *The Kaohsiung journal of medical sciences*, 25(3):116–125, 2009.

[191] Nakashima, Y., Inukai, K., Imai, K., Ikegami, Y., Awata, T., and Katayama, S. Involvement of low adiponectin levels in impaired glucose tolerance. *Metabolism*, 57(10):1350–1354, 2008.

[192] Turpeinen, A., Haffner, S., Louheranta, A., Niskanen, L., Miettinen, H., and Uusitupa, M. Serum leptin in subjects with impaired glucose tolerance in relation to insulin sensitivity and first-phase insulin response. *International journal of obesity and related metabolic disorders: journal of the International Association for the Study of Obesity*, 21(4):284–287, 1997.

[193] Rask, E., Olsson, T., Söderberg, S., Johnson, O., Seckl, J., Holst, J. J., and Ahrén, B. Impaired incretin response after a mixed meal is associated with insulin resistance in nondiabetic men. *Diabetes Care*, 24(9):1640–1645, 2001.

[194] Verdich, C., Toubro, S., Buemann, B., Lysgård, M. J., Juul, H. J., and Astrup, A. The role of postprandial releases of insulin and incretin hormones in meal-induced satiety–effect of obesity and weight reduction. *International journal of obesity and related metabolic disorders: journal of the International Association for the Study of Obesity*, 25(8):1206–1214, 2001.

[195] Ranganath, L., Norris, F., Morgan, L., Wright, J., and Marks, V. Inhibition of carbohydrate-mediated glucagon-like peptide-i (7-36) amide secretion by circulating non-esterified fatty acids. *Clinical Science*, 96:335–342, 1999.

[196] Ting, C., Rosenberg, M., Snow, C., Samuelson, L., and Meisler, M. Endogenous retroviral sequences are required for tissue-specific expression of a human salivary amylase gene. *Genes & development*, 6(8):1457–1465, 1992.

[197] Holst, J. J., Knop, F. K., Vilsbøll, T., Krarup, T., and Madsbad, S. Loss of incretin effect is a specific, important, and early characteristic of type 2 diabetes. *Diabetes Care*, 34(Supplement 2):S251–S257, 2011.

[198] Kang, H. M., Zaitlen, N. A., Wade, C. M., Kirby, A., Heckerman, D., Daly, M. J., and Eskin, E. Efficient control of population structure in model organism association mapping. *Genetics*, 178(3):1709–1723, 2008.

[199] Kang, H. M., Sul, J. H., Service, S. K., Zaitlen, N. A., Kong, S.-y., Freimer, N. B., Sabatti, C., Eskin, E., et al. Variance component model to account for sample structure in genome-wide association studies. *Nature genetics*, 42(4):348–354, 2010.

[200] Lipka, A. E., Tian, F., Wang, Q., Peiffer, J., Li, M., Bradbury, P. J., Gore, M. A., Buckler, E. S., and Zhang, Z. Gapit: genome association and prediction integrated tool. *Bioinformatics*, 28(18):2397–2399, 2012.

[201] Fisher, R. A., et al. The correlation between relatives on the supposition of mendelian inheritance. *Transactions of the Royal Society of Edinburgh*, 52:399–433, 1918.

[202] Henderson, C., et al. Applications of linear models in animal breeding. *Applications of linear models in animal breeding.*, 1984.

[203] Yu, J., Pressoir, G., Briggs, W. H., Bi, I. V., Yamasaki, M., Doebley, J. F., McMullen, M. D., Gaut, B. S., Nielsen, D. M., Holland, J. B., et al. A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. *Nature genetics*, 38(2):203–208, 2006.

[204] Vilhjálmsson, B. J., and Nordborg, M. The nature of confounding in genome-wide association studies. *Nature Reviews Genetics*, 14(1):1–2, 2013.

[205] Segura, V., Vilhjálmsson, B. J., Platt, A., Korte, A., Seren, Ü., Long, Q., and Nordborg, M. An efficient multi-locus mixed-model approach for genome-wide association studies in structured populations. *Nature genetics*, 44(7):825–830, 2012.

[206] Lippert, C., Listgarten, J., Liu, Y., Kadie, C. M., Davidson, R. I., and Heckerman, D. Fast linear mixed models for genome-wide association studies. *Nature Methods*, 8(10):833–835, 2011.

[207] Zhou, X., Carbonetto, P., and Stephens, M. Polygenic modeling with bayesian sparse linear mixed models. *PLoS Genet*, 9(2):e1003264, 2013.

[208] Widmer, C., Lippert, C., Weissbrod, O., Fusi, N., Kadie, C., Davidson, R., Listgarten, J., and Heckerman, D. Further improvements to linear mixed models for genome-wide association studies. *Scientific reports*, 4, 2014.

[209] Wang, Q., Tian, F., Pan, Y., Buckler, E. S., and Zhang, Z. A super powerful method for genome wide association study. 2014.

[210] Szydlowski, M., and Paczyńska, P. Qtlmas 2010: simulated dataset. In *BMC proceedings*, volume 5, page S3. BioMed Central Ltd, 2011.

[211] Browning, S. R., and Browning, B. L. Rapid and accurate haplotype phasing and missing-data inference for whole-genome association studies by use of localized haplotype clustering. *The American Journal of Human Genetics*, 81(5):1084–1097, 2007.

[212] Browning, B. L., and Browning, S. R. Improving the accuracy and efficiency of identity-by-descent detection in population data. *Genetics*, 194(2):459–471, 2013.

[213] Zhou, X., and Stephens, M. Genome-wide efficient mixed-model analysis for association studies. *Nature genetics*, 44(7):821–824, 2012.

[214] Hoffman, G. E. Correcting for population structure and kinship using the linear mixed model: theory and extensions. 2013.

[215] Schaid, D. J. Genomic similarity and kernel methods i: advancements by building on mathematical and statistical foundations. *Human heredity*, 70(2):109–131, 2010.

[216] Listgarten, J., Lippert, C., Kadie, C. M., Davidson, R. I., Eskin, E., and Heckerman, D. Improved linear mixed models for genome-wide association studies. *Nature methods*, 9(6):525–526, 2012.

[217] Lippert, C., Xiang, J., Horta, D., Widmer, C., Kadie, C., Heckerman, D., and Listgarten, J. Greater power and computational efficiency for kernel-based association testing of sets of genetic variants. *Bioinformatics*, page btu504, 2014.

[218] Allen, H. L., Estrada, K., Lettre, G., Berndt, S. I., Weedon, M. N., Rivadeneira, F., Willer, C. J., Jackson, A. U., Vedantam, S., Raychaudhuri, S., et al. Hundreds of variants clustered in genomic loci and biological pathways affect human height. *Nature*, 467(7317):832–838, 2010.

[219] Frischknecht, M., Jagannathan, V., Plattet, P., Neuditschko, M., Signer-Hasler, H., Bachmann, I., Pacholewska, A., Drögemüller, C., Dietschi, E., Flury, C., et al. A non-synonymous hmga2 variant decreases height in shetland ponies and other small horses. *PloS one*, 10(10):e0140749, 2015.

[220] Makvandi-Nejad, S., Hoffman, G. E., Allen, J. J., Chu, E., Gu, E., Chandler, A. M., Loredo, A. I., Bellone, R. R., Mezey, J. G., Brooks, S. A., et al. Four loci explain 83% of size variation in the horse. *PLoS One*, 7(7):e39929, 2012.

[221] Signer-Hasler, H., Flury, C., Haase, B., Burger, D., Simianer, H., Leeb, T., and Rieder, S. A genome-wide association study reveals loci influencing height and other conformation traits in horses. *PLoS One*, 7(5):e37282, 2012.

[222] Pryce, J. E., Hayes, B. J., Bolormaa, S., and Goddard, M. E. Polymorphic regions affecting human height also control stature in cattle. *Genetics*, 187(3):981–984, 2011.

[223] Korte, A., Vilhjálmsson, B. J., Segura, V., Platt, A., Long, Q., and Nordborg, M. A mixed-model approach for genome-wide association studies of correlated traits in structured populations. *Nature genetics*, 44(9):1066–1071, 2012.

[224] Listgarten, J., Lippert, C., and Heckerman, D. Fast-lmm-select for addressing confounding from spatial structure and rare variants. *Nature Genetics*, 45(5):470–471, 2013.

[225] Yang, J., Zaitlen, N. A., Goddard, M. E., Visscher, P. M., and Price, A. L. Advantages and pitfalls in the application of mixed-model association methods. *Nature genetics*, 46(2):100–106, 2014.

[226] Zhang, Z., Ersoz, E., Lai, C.-Q., Todhunter, R. J., Tiwari, H. K., Gore, M. A., Bradbury, P. J., Yu, J., Arnett, D. K., Ordovas, J. M., et al. Mixed linear model approach adapted for genome-wide association studies. *Nature genetics*, 42(4):355–360, 2010.

[227] Lippert, C., Quon, G., Kang, E. Y., Kadie, C. M., Listgarten, J., and Heckerman, D. The benefits of selecting phenotype-specific variants for applications of mixed models in genomics. *Scientific reports*, 3, 2013.

[228] Habier, D., Fernando, R. L., Kizilkaya, K., and Garrick, D. J. Extension of the bayesian alphabet for genomic selection. *BMC bioinformatics*, 12(1):186, 2011.

[229] Loh, P.-R., Tucker, G., Bulik-Sullivan, B. K., Vilhjalmsson, B. J., Finucane, H. K., Salem, R. M., Chasman, D. I., Ridker, P. M., Neale, B. M., Berger, B., et al. Efficient bayesian mixed-model analysis increases association power in large cohorts. *Nature genetics*, 2015.

[230] Schröder, W., Klostermann, A., Stock, K., and Distl, O. A genome-wide association study for quantitative trait loci of show-jumping in hanoverian warmblood horses. *Animal genetics*, 43(4):392–400, 2012.

[231] Teyssèdre, S., Dupuis, M., Guérin, G., Schibler, L., Denoix, J., Elsen, J. M., and Ricard, A. Genome-wide association studies for osteochondrosis in french trotter horses. *Journal of animal science*, 90(1):45–53, 2012.

[232] Lykkjen, S., Dolvik, N., McCue, M., Rendahl, A., Mickelson, J., and Roed, K. Genome-wide association analysis of osteochondrosis of the tibiotarsal joint in norwegian standardbred trotters. *Animal genetics*, 41(s2):111–120, 2010.

[233] Cody, J. D., Heard, P., and Hale, D. Identification of two novel chromosome regions associated with isolated growth hormone deficiency. *Journal of Pediatric Endocrinology and Metabolism*, 23(11):1159–1164, 2010.

[234] Asa, S. L., DiGiovanni, R., Jiang, J., Ward, M. L., Loesch, K., Yamada, S., Sano, T., Yoshimoto, K., Frank, S. J., and Ezzat, S. A growth hormone receptor mutation impairs growth hormone autofeedback signaling in pituitary tumors. *Cancer research*, 67(15):7505–7511, 2007.

[235] Valverde, P., Healy, E., Jackson, I., Rees, J. L., and Thody, A. J. Variants of the melanocyte–stimulating hormone receptor gene are associated with red hair and fair skin in humans. *Nature genetics*, 11(3):328–330, 1995.

[236] Andersson, L. Melanocortin receptor variants with phenotypic effects in horse, pig, and chicken. *Annals of the New York Academy of Sciences*, 994(1):313–318, 2003.

[237] Wilson, C. L., Liu, W., Yang, J. J., Kang, G., Ojha, R. P., Neale, G. A., Srivastava, D. K., Gurney, J. G., Hudson, M. M., Robison, L. L., et al. Genetic and clinical factors associated with obesity among adult survivors of childhood cancer: A report from the st. jude lifetime cohort. *Cancer*, 2015.

[238] Reiner, A. P., Beleza, S., Franceschini, N., Auer, P. L., Robinson, J. G., Kooperberg, C., Peters, U., and Tang, H. Genome-wide association and population genetic analysis of c-reactive protein in african american and hispanic american women. *The American Journal of Human Genetics*, 91(3):502–512, 2012.

[239] Fox, C. S., Heard-Costa, N., Cupples, L. A., Dupuis, J., Vasan, R. S., and Atwood, L. D. Genome-wide association to body mass index and waist circumference: the framingham heart study 100k project. *BMC medical genetics*, 8(Suppl 1):S18, 2007.

[240] Kathiresan, S., Manning, A. K., Demissie, S., D'agostino, R. B., Surti, A., Guiducci, C., Gianniny, L., Burtt, N. P., Melander, O., Orho-Melander, M., et al. A genome-wide association study for blood lipid phenotypes in the framingham heart study. *BMC medical genetics*, 8(Suppl 1):S17, 2007.

[241] Levy, D., Larson, M. G., Benjamin, E. J., Newton-Cheh, C., Wang, T. J., Hwang, S.-J., Vasan, R. S., and Mitchell, G. F. Framingham heart study 100k project: genome-wide associations for blood pressure and arterial stiffness. *BMC medical genetics*, 8(Suppl 1):S3, 2007.

[242] Surakka, I., Isaacs, A., Karssen, L. C., Laurila, P., Middelberg, R. P., Tikkanen, E., Ried, J. S., Lamina, C., Mangino, M., Igl, W., et al. A genome-wide screen for interactions reveals a new locus on 4p15 modifying the effect of waist-to-hip ratio on total cholesterol. *PLoS Genet*, 7(10):e1002333, 2011.

[243] Cornelis, M. C., Monda, K. L., Yu, K., Paynter, N., Azzato, E. M., Bennett, S. N., Berndt, S. I., Boerwinkle, E., Chanock, S., Chatterjee, N., et al. Genome-wide meta-analysis identifies regions on 7p21 (ahr) and 15q24 (cyp1a2) as determinants of habitual caffeine consumption. *PLoS Genet*, 7(4):e1002033, 2011.

[244] Newton-Cheh, C., Guo, C.-Y., Wang, T. J., O'donnell, C. J., Levy, D., and Larson, M. G. Genome-wide association study of electrocardiographic and heart rate variability traits: the framingham heart study. *BMC medical genetics*, 8(Suppl 1):S7, 2007.

[245] Teslovich, T. M., Musunuru, K., Smith, A. V., Edmondson, A. C., Stylianou, I. M., Koseki, M., Pirruccello, J. P., Ripatti, S., Chasman, D. I., Willer, C. J., et al. Biological, clinical and population relevance of 95 loci for blood lipids. *Nature*, 466(7307):707–713, 2010.

[246] Consortium, G. L. G., et al. Discovery and refinement of loci associated with lipid levels. *Nature genetics*, 45(11):1274–1283, 2013.

[247] Comuzzie, A. G., Cole, S. A., Laston, S. L., Voruganti, V. S., Haack, K., Gibbs, R. A., and Butte, N. F. Novel genetic loci identified for the pathophysiology of childhood obesity in the hispanic population. 2012.

[248] Meigs, J. B., Manning, A. K., Fox, C. S., Florez, J. C., Liu, C., Cupples, L. A., and Dupuis, J. Genome-wide association with diabetes-related traits in the framingham heart study. *BMC medical genetics*, 8(Suppl 1):S16, 2007.

[249] Dastani, Z., Hivert, M.-F., Timpson, N., Perry, J. R., Yuan, X., Scott, R. A., Henneman, P., Heid, I. M., Kizer, J. R., Lyytikäinen, L.-P., et al. Novel loci for adiponectin levels and their influence on type 2 diabetes and metabolic traits: a multi-ethnic meta-analysis of 45,891 individuals. *PLoS Genet*, 8(3):e1002607, 2012.

[250] Heid, I. M., Jackson, A. U., Randall, J. C., Winkler, T. W., Qi, L., Steinthorsdottir, V., Thorleifsson, G., Zillikens, M. C., Speliotes, E. K., Mägi, R., et al. Meta-analysis identifies 13 new loci associated with waist-hip ratio and reveals sexual dimorphism in the genetic basis of fat distribution. *Nature genetics*, 42(11):949–960, 2010.

[251] Jung, M. Y., Kim, B. S., Kim, Y. J., Koh, I. S., and Chung, J.-H. Assessment of relationship between fyn-related kinase gene polymorphisms and overweight/obesity in korean population. *The Korean Journal of Physiology & Pharmacology*, 12(2):83–87, 2008.

[252] Speliotes, E. K., Willer, C. J., Berndt, S. I., Monda, K. L., Thorleifsson, G., Jackson, A. U., Allen, H. L., Lindgren, C. M., Luan, J., Mägi, R., et al. Association analyses of 249,796 individuals reveal 18 new loci associated with body mass index. *Nature genetics*, 42(11):937–948, 2010.

[253] Wang, K., Li, W.-D., Zhang, C. K., Wang, Z., Glessner, J. T., Grant, S., Zhao, H., Hakonarson, H., and Price, R. A. A genome-wide association study on obesity and obesity-related traits. *PloS one*, 6(4):e18939, 2011.

[254] Heard-Costa, N. L., Zillikens, M. C., Monda, K. L., Johansson, A., Harris, T. B., Fu, M., Haritunians, T., Feitosa, M. F., Aspelund, T., Eiriksdottir, G., et al. Nrxn3 is a novel locus for waist circumference: a genome-wide association study from the charge consortium. *PLoS Genet*, 5(6):e1000539, 2009.

[255] Flicek, P., Amode, M. R., Barrell, D., Beal, K., Billis, K., Brent, S., Carvalho-Silva, D., Clapham, P., Coates, G., Fitzgerald, S., et al. Ensembl 2014. *Nucleic acids research*, page gkt1196, 2013.

[256] NOVAK, U., WILKS, A., BUELL, G., and McEWEN, S. Identical mrna for preproglucagon in pancreas and gut. *European Journal of Biochemistry*, 164(3):553–558, 1987.

Appendix A

Supplemental Material: Chapter 2 Re-defining the equine metabolic syndrome phenotype

Figure A.1: Initial online survey for Equine Metabolic Syndrome.

CONTACT INFORMATION
Your Name: (fill in the blank)
Phone Number: (fill in the blank)
Your e-mail: (fill in the blank)
Horse's Call (Barn) Name: (fill in the blank)
Horse's Registered Name: (fill in the blank)

VETERINARIAN'S CONTACT INFORMATION
Veterinarian Name: (fill in the blank)
Clinic Name: (fill in the blank)
Clinic Address: (fill in the blank)
Clinic Phone: (fill in the blank)
Veterinarian/Clinic e-mail: (fill in the blank)

SIGNALMENT
1. Horse's (approximate) year of birth? (drop down choices: 1975-2015)
2. What breed is the horse? (drop down choices: 56 different breeds and "other")
If you chose other please provide breed(s) (fill in the blank)
3. What is the horse's gender? (drop down choices: male intact/ male castrated/ female intact/ female spayed)
4. Would you describe your horse as an "easy keeper" (tends to gain weight easily)? (drop down choices: Yes/No/Not Sure)

LAMINITIS STATUS
5. Has your horse been previously diagnosed with laminitis? (drop down choices: Yes/No/Not Sure)
6. Please fill in the blank with any additional medical conditions occurring at the time of laminitis if "yes" was selected for the previous question? (fill in the blank)

Figure A.1 *Continued from previous page*: Initial online survey for Equine Metabolic Syndrome.

PHENOTYPE

7. Is your horse overweight? (drop down choices: Yes/No/Not Sure)
8. Is your horse underweight? (drop down choices: Yes/No/Not Sure)
9. Does your horse have a thick, cresty neck? (drop down choices: Yes/No/Not Sure)
10. Does your horse have fat pads on its rump? (drop down choices: Yes/No/Not Sure)
11. Does your horse have hair coat that sheds later in the year? (drop down choices: Yes/No/Not Sure)
12. Does your horse have thick or long curly hair coat? (drop down choices: Yes/No/Not Sure)
13. Does your horse urinate more than normal? (drop down choices: Yes/No/Not Sure)
14. Does your horse drink more than normal? (drop down choices: Yes/No/Not Sure)

PREVIOUS LAB WORK

15. Has your horse had any of the following lab work done?
Insulin Level: (drop down choices: Yes, High /Yes, Normal /Yes, Low /Not performed)
Glucose Level: (drop down choices: Yes, High /Yes, Normal /Yes, low /Not performed)
Dexamethasone Suppression Test: (drop down choices: Yes, Normal Suppression /Yes, Abnormal Suppression /Not Performed)

OTHER INFORMATION

16. How many other horses are kept on the same property as your horse? (fill in the blank)
17. How many other horses kept on the same property have experienced laminitis in the past? (fill in the blank)
18. How many other horses on the same property are overweight or have been diagnosed with equine metabolic syndrome? (fill in the blank)
19. May we contact you for more information? (drop down choices: Yes/No)

Figure A.2: Second online survey for Equine Metabolic Syndrome.

Based on results from the initial survey, owners were given a link to access a second online survey to gather information about a potential metabolic syndrome horse and a second suitable control on the property.

CONTACT INFORMATION
Owner Last Name: (fill in the blank)
Owner First Name: (fill in the blank)
Owner Email: (fill in the blank)
Horse's Call (Barn) Name: (fill in the blank)
Horse's Registered Name: (fill in the blank)
Owner Address: (fill in the blank)
City: (fill in the blank)
State: (fill in the blank)
Zip: (fill in the blank)
*contact information was used to link first and second survey results

HORSE MEASUREMENTS*
*online description and diagrams of requested measurements were available on the survey website
1. What is your horse?s body condition score on a scale of 1-9? (drop down choices 1-9 in 0.5 unit increments / not sure)
2. Horse Photos: horse owners were asked to upload digital photos for 4 standard views of horses
3. What is your horse's body length? (in inches) (fill in the blank)
4. What is your horse's neck circumference measurement? (in inches) (fill in the blank)
5. What is your horse's girth measurement? (in inches) (fill in the blank)
6. What is your horse's weight? (in pounds) (fill in the blank)
7. How long has the horse been considered overweight? Not applicable (horse is not overweight) /less than 1 year /1 year /2 years /three years /4-6 years /more than 6 years /not sure
8. How was this weight determined? Estimate /scale /weight tape /not sure
9. What is your horse's height at the withers (measured in inches)? fill in the blank

Figure A.2 *Continued from previous page*: Second online survey for Equine Metabolic Syndrome.

PREVIOUS DIAGNOSIS

10. Has the horse been diagnosed with metabolic syndrome? (drop down choices: Yes/No/Not Sure)

11. Has the horse ever been diagnosed with Cushings disease (pituitary pars intermedia dysfunction, PPID)? (drop down choices: Yes/No/Not Sure)

PREVIOUS LABORATORY DATA

Please work with your veterinarian to enter the horse's values for any of the following laboratory data if available (please provide units of measure):

12. Insulin concentration: fill in the blank

13. When was the sample for insulin concentration taken? drop down choices:

- not applicable (test not performed)
- the horse was not fasted
- after a complete fast (no hay or grain) of greater than 6 hours
- after a complete fast (no hay or grain) of greater than 10 hours
- after a partial fast (hay but no grain) of greater than 6 hours
- after a partial fast (hay but no grain) of greater than 10 hours
- not sure

14. The insulin concentration was _____compared to normal values. (drop down choices: not applicable (test not performed) /high low /normal /not sure)

15. Glucose concentration: fill in the blank

16. When was the sample for glucose concentration taken? drop down choices:

- not applicable (test not performed)
- the horse was not fasted
- after a complete fast (no hay or grain) of greater than 6 hours
- after a complete fast (no hay or grain) of greater than 10 hours
- after a partial fast (hay but no grain) of greater than 6 hours
- after a partial fast (hay but no grain) of greater than 10 hours
- not sure

17. The glucose concentration was_____compared to normal values. (drop down choices: not applicable (test not performed) /high low /normal /not sure)

Figure A.2 *Continued from previous page*: Second online survey for Equine Metabolic Syndrome.

PREVIOUS LABORATORY DATA *continued*

18. Triglyceride concentration: fill in the blank

19. When was the sample for triglyceride concentration taken? drop down choices:

- not applicable (test not performed)
- the horse was not fasted
- after a complete fast (no hay or grain) of greater than 6 hours
- after a complete fast (no hay or grain) of greater than 10 hours
- after a partial fast (hay but no grain) of greater than 6 hours
- after a partial fast (hay but no grain) of greater than 10 hours
- not sure

20. The triglyceride concentration was _____compared to normal values. (drop down choices: not applicable (test not performed) /high low /normal /not sure)

21. Nonesterified fatty acids (NEFA) concentration: fill in the blank

22. When was the sample for NEFA concentration taken? drop down choices:

- not applicable (test not performed)
- the horse was not fasted
- after a complete fast (no hay or grain) of greater than 6 hours
- after a complete fast (no hay or grain) of greater than 10 hours
- after a partial fast (hay but no grain) of greater than 6 hours
- after a partial fast (hay but no grain) of greater than 10 hours
- not sure

23. The NEFA concentration was _____compared to normal values. (drop down choices: not applicable (test not performed) /high low /normal /not sure)

24. ACTH concentration: fill in the blank

25. What time of day was the sample for ACTH concentration taken? drop down choices:

- not applicable (test not performed)
- early morning
- mid-day
- evening
- not sure

26. What month was the ACTH concentration taken? drop down (January-December)

27. The ACTH concentration was _____compared to normal values. (drop down choices: not applicable (test not performed) /high low /normal /not sure)

Figure A.2 *Continued from previous page*: Second online survey for Equine Metabolic Syndrome.

PREVIOUS LABORATORY DATA *continued*

28. Cortisol concentration: fill in the blank

29. What time of day was the cortisol concentration measured? drop down choices:

- not applicable (test not performed)
- early morning
- mid-day
- evening
- not sure

30. What month was the cortisol concentration taken? Drop down (January-December)

31. The cortisol concentration was _____compared to normal values. (drop down choices: not applicable (test not performed) /high low /normal /not sure)

LAMINITIS STATUS

Please answer the following questions regarding the horse's laminitis (if applicable):

32. Was the horse experiencing laminitis when any of the above tests were performed? (drop down choices: Yes/No/Not Sure)

33. Has the horse had any of the following endocrine testing done? (please choose all that apply)

- 24 hour dexamathasone suppression test
- overnight dexamathasone suppression test
- ACTH stimulation test
- combined dexamathasone suppression and ACTH stimulation test
- thyrotropin-releasing hormone stimulation test
- glucose tolerance test
- insulin tolerance test
- not sure
- Other:_____

34. Has your horse been previously diagnosed with laminitis? (drop down choices: Yes/No/Not Sure)

35. If so, at what age did your horse first develop laminitis? fill in the blank

Figure A.2 *Continued from previous page*: Second online survey for Equine Metabolic Syndrome.

LAMINITIS STATUS *continued*

36. In what year did the most recent laminitis episode occur? drop down choices:

- not applicable (horse has not had laminitis)
- 2010
- 2009
- 2008
- 2007
- 2006 2005 2004 .... 1991 1990 1989
- not sure

37. In what month did the most recent laminitis episode occur? (drop down choices: not applicable (horse has not had laminitis) / months January-December / not sure)

38. Which limbs are affected? drop down choices:

- not applicable (horse has not had laminitis)
- forelimbs
- hindlimbs
- all four limbs
- not sure

39. How have the horse's symptoms changed over time? drop down choices:

- improved
- worsened
- stayed the same
- not applicable
- not sure

40. What treatments has the horse received for laminitis? (Pick all that apply)

- no treatment
- corrective shoeing
- phenylbutazone (bute)
- flunixin meglumine (Banamine)
- not sure

41. Do you think your horse's laminitis problems are related to grazing on lush green pastures? (drop down choices: not applicable (horse has not had laminitis) / Yes /No /not sure)

42. How frequently does you horse have bouts of laminitis? drop down choices:

- not applicable (horse has not had laminitis)
- once every 2-3 years
- once per year
- 2 times per year
- 3 times per year
- more than 3 time per year
- occurs year long
- not sure

Figure A.2 *Continued from previous page*: Second online survey for Equine Metabolic Syndrome.

PREVIOUS TREATMENTS

43. What treatments has the horse received for its metabolic syndrome/obesity? (Choose all that apply) (drop down)

- chromium supplementation
- magnesium supplementation
- thyroid supplementation
- metformin
- dietary change
- management change
- not sure Other:_____

44. Has the diet been changed? (drop down choices: decreased /increased /not changed /not sure)

45. Has the amount of turnout been changed? (drop down choices: decreased /increased /not changed /not sure)

46. Has the amount of exercise changed? (drop down choices: decreased /increased /not changed /not sure)

DIET AND MANAGEMENT

Please answer the following questions about the horse's diet and management when **signs consistent with metabolic syndrome** were first apparent:

47. How much hay did the horse receive per day? (drop down)

- less than 10 lbs
- 10-14 lbs
- 15-19 lbs
- more than 20 lbs
- not sure

48. What type of hay was fed? (drop down)

- grass hay
- legume (alfalfa)
- mix (predominantly grass)
- mix (predominantly alfalfa)
- not sure

Figure A.2 *Continued from previous page*: Second online survey for Equine Metabolic Syndrome.

DIET AND MANAGEMENT*continued*

49. How much grain did the horse receive per day? (drop down)

- none
- less than 1 lb
- 1-2 lbs
- 3-4 lbs
- more than 5 lbs
- not sure

50. What supplements did the horse receive? (Choose all that apply)

- No supplements
- Vitamin E
- Fat/oil
- Selenium
- Mineral block
- Salt block
- Electrolyte
- Joint supplement
- not sure
- Other:_____

51. How much time did the horse spend in a stall on average? (drop down)

- 0 - 1 hour/day
- 1 - 6 hours/day
- 6 - 12 hours/day
- 12+ hours/day
- not sure

52. How much riding/forced exercise did the horse get per week? (drop down)

- none
- less than 30 minutes
- 30 min to 60 min
- 60 min to 120 min
- 120 min to 180 min
- 180 min to 240 min
- more than 240 min
- not sure

53. How much grass pasture turnout did the horse receive? (drop down)

- none
- less than 1 hour per day
- 1-3 hours per day
- 3-6 hours per day
- 6-9 hours per day
- more than 9 hours per day
- 100% turnout
- not sure

Figure A.2 *Continued from previous page*: Second online survey for Equine Metabolic Syndrome.

DIET AND MANAGEMENT*continued*
Please answer the following questions about the horse's **current** diet and management:

54. How much hay does the horse receive per day? (drop down)
- less than 10 lbs
- 10-14 lbs
- 15-19 lbs
- more than 20 lbs
- not sure

55. What type of hay is fed? (drop down)
- grass hay
- legume (alfalfa)
- mix (predominantly grass)
- mix (predominantly alfalfa)
- not sure

56. How much grain does the horse receive per day? (drop down)
- none
- less than 1 lb
- 1-2 lbs
- 3-4 lbs
- more than 5 lbs
- not sure

57. What supplements does the horse receive? (Choose all that apply) (drop down)
- No supplements
- Vitamin E
- Fat/oil
- Selenium
- Mineral block
- Salt block
- Electrolyte
- Joint supplement
- not sure
- Other: _____

58. How much time does the horse spend in a stall on average? (drop down)
- 0 - 1 hour/day
- 1 - 6 hours/day
- 6 - 12 hours/day
- 12+ hours/day
- not sure

Figure A.2 *Continued from previous page*: Second online survey for Equine Metabolic Syndrome.

DIET AND MANAGEMENT*continued*
Please answer the following questions about the horse's **current** diet and management:

59. How much riding/forced exercise does the horse get per week? (drop down)

- none
- less than 30 minutes
- 30 min to 60 min
- 60 min to 120 min
- 120 min to 180 min
- 180 min to 240 min
- more than 240 min
- not sure

60. How much grass pasture turnout does the horse receive? (drop down)

- none
- less than 1 hour per day
- 1-3 hours per day
- 3-6 hours per day
- 6-9 hours per day
- more than 9 hours per day
- 100% turnout
- not sure

61. What other chronic health problems does the horse have? fill in the blank/short answer

62. What is the horse used for? (Choose all that apply)

- Pleasure/trail riding
- Lounging
- Driving
- Dressage
- Eventing
- Hunter/jumper
- Racing
- Roping
- Cutting
- Barrel racing
- Halter
- Other: _____

63. Do any related horses have similar symptoms? (drop down)

- Yes
- No

Figure A.2 *Continued from previous page*: Second online survey for Equine Metabolic Syndrome.

CONTROL HORSE INFORMATION

Please answer the remaining questions for a second horse on the same property not suspected of having EMS (to serve as "control")

Owner name (required): (fill in the blank)
Address: (fill in the blank)
phone: (fill in the blank)
e-mail (required): (fill in the blank)
Horse's call (barn) name: (fill in the blank)
Horse's Registered name (if registered): (fill in the blank)
Control Horse Photos: (submitted as above)

CONTROL HORSE SIGNALMENT

1. Horse's year of birth? (drop down choices: 1975-2015)
2. What breed is the horse? (drop down choices: 56 different breeds and "other") If you chose other please provide breed(s) (fill in the blank)
3. What is the horse's gender? (drop down choices: male intact/ male castrated/ female intact/ female spayed)
4. Would you describe your horse as an "easy keeper"(tends to gain weight easily)? (drop down choices: Yes/No/Not Sure)

CONTROL HORSE PHENOTYPE

Which of the following descriptions apply to your horse (control) now or in the past?
5. Overweight (drop down choices: Yes/No/Not Sure)
6. Underweight (drop down choices: Yes/No/Not Sure)
7. Thick, cresty neck (drop down choices: Yes/No/Not Sure)
8. Fat pads on rump (drop down choices: Yes/No/Not Sure)
9. Hair coat that sheds later in year (drop down choices: Yes/No/Not Sure)
10. Thick or long curly hair coat (drop down choices: Yes/No/Not Sure)
11. Urinates more than normal (drop down choices: Yes/No/Not Sure)
12. Drinks more than normal (drop down choices: Yes/No/Not Sure)
13. Sweats more than normal (drop down choices: Yes/No/Not Sure)

Figure A.2 *Continued from previous page*: Second online survey for Equine Metabolic Syndrome.

CONTROL HORSE MEASUREMENTS*

*online description and diagrams of requested measurements were available on the survey website

14. What is your horse's body condition score on a scale of 1-9? (drop down choices 1-9 in 0.5 unit increments / not sure)

15. Horse Photos: horse owners were asked to upload digital photos for 4 standard views of horses

16. What is your horse's body length? (in inches) (fill in the blank)

17. What is your horse's neck circumference measurement? (in inches) (fill in the blank)

18. What is your horse's girth measurement? (in inches) (fill in the blank)

19. What is your horse's weight? (in pounds) (fill in the blank)

20. How long has the horse been considered overweight? Not applicable (horse is not overweight) /less than 1 year /1 year /2 years /three years /4-6 years /more than 6 years /not sure

21. How was this weight determined? Estimate /scale /weight tape /not sure

22. What is your horse's height at the withers (measured in inches)? fill in the blank

CONTROL HORSE PREVIOUS DIAGNOSIS

23. Has the horse been diagnosed with metabolic syndrome? (drop down choices: Yes/No/Not Sure)

24. Has the horse ever been diagnosed with Cushings disease (pituitary pars intermedia dysfunction, PPID)? (drop down choices: Yes/No/Not Sure)

25. Has your horse been previously diagnosed with laminitis? (drop down choices: Yes/No/Not Sure)

Figure A.2 *Continued from previous page*: Second online survey for Equine Metabolic Syndrome.

CONTROL HORSE DIET AND MANAGEMENT

Please answer the following questions about the control horse's current diet and management:

26. How much hay does the horse receive per day? (drop down)

- less than 10 lbs
- 10-14 lbs
- 15-19 lbs
- more than 20 lbs
- not sure

27. What type of hay is fed? (drop down)

- grass hay
- legume (alfalfa)
- mix (predominantly grass)
- mix (predominantly alfalfa)
- not sure

28. How much grain does the horse receive per day? (drop down)

- none
- less than 1 lb
- 1-2 lbs
- 3-4 lbs
- more than 5 lbs
- not sure

29. What supplements does the horse receive? (Choose all that apply) (drop down)

- No supplements
- Vitamin E
- Fat/oil
- Selenium
- Mineral block
- Salt block
- Electrolyte
- Joint supplement
- not sure
- Other: _____

Figure A.2 *Continued from previous page*: Second online survey for Equine Metabolic Syndrome.

CONTROL HORSE DIET AND MANAGEMENT *continued*

30. How much time does the horse spend in a stall on average? (drop down)

- 0 - 1 hour/day
- 1 - 6 hours/day
- 6 - 12 hours/day
- 12+ hours/day
- not sure

31. How much riding/forced exercise does the horse get per week? (drop down)

- none
- less than 30 minutes
- 30 min to 60 min
- 60 min to 120 min
- 120 min to 180 min
- 180 min to 240 min
- more than 240 min
- not sure

32. How much grass pasture turnout does the horse receive? (drop down)

- none
- less than 1 hour per day
- 1-3 hours per day
- 3-6 hours per day
- 6-9 hours per day
- more than 9 hours per day
- 100% turnout
- not sure

33. What chronic health problems does the horse have? fill in the blank/short answer

34. What is the horse used for? (Choose all that apply)

- Pleasure/trail riding
- Lounging
- Driving
- Dressage
- Eventing
- Hunter/jumper
- Racing
- Roping
- Cutting
- Barrel racing
- Halter
- Other: _____

35. May we contact you for more information? (drop down choices Yes / No)

Table A.1: Metabolic phenotype descriptive summary statistics for all levels of categorical explanatory variables. Mean (SD), median, number of individuals.

Categorical explanatory variable	NH	GH
Non-obese, no prior laminitis	0.65(0.05),0.64, n=335	1.20(0.06),1.21, n=339
Non-obese, prior laminitis	0.66(0.05),0.67, n=95	1.20(0.09),1.22, n=95
Obese, no prior laminitis	0.68(0.06),0.68, n=110	1.27(0.06),1.27, n=111
Obese, prior laminitis	0.68(0.06),0.68, n=63	1.26(0.08),1.26, n=63
Morgan	0.67(0.05),0.67, n=288	1.23(0.06),1.23, n=288
Arabian	0.63(0.05),0.63, n=62	1.19(0.06),1.20, n=62
Other high risk breeds	0.64(0.06),0.63, n=45	1.20(0.08),1.21, n=45
Other low risk breeds	0.61(0.05),0.62, n=16	1.19(0.08),1.17, n=16
Pony	0.67(0.05),0.66, n=96	1.25(0.07),1.25, n=96
Quarter Horse	0.65(0.07),0.65, n=53	1.24(0.07),1.24, n=57
Tennessee Walking Horse	0.64(0.04),0.64, n=43	1.18(0.08),1.20, n=44
Gelding	0.65(0.05),0.65, n=206	1.21(0.07),1.21, n=209
Mare	0.66(0.05),0.66, n=365	1.23(0.07),1.24, n=367
Stallion	0.69(0.05),0.69, n=32	1.18(0.05),1.17, n=32
Owner submitted sample	0.67(0.05),0.67, n=333	1.23(0.07),1.23, n=333
Researcher collected sample	0.64(0.06),0.64, n=270	1.21(0.07),1.21, n=275
No L-thyroxine supplementation	0.66(0.05),0.66, n=579	1.22(0.07),1.23, n=584
Yes L-thyroxine supplementation	0.64(0.06),0.64, n=24	1.19(0.06),1.17, n=24
DEC collection	0.66(0.04),0.66, n=58	1.23(0.06),1.23, n=58
JAN collection	0.63(0.05),0.64, n=8	1.23(0.03),1.24, n=7
FEB collection	0.68(0.05),0.67, n=35	1.25(0.06),1.25, n=39
MAR collection	0.65(0.05),0.66, n=86	1.22(0.05),1.21, n=86
APR collection	0.66(0.05),0.66, n=69	1.23(0.07),1.24, n=69
MAY collection	0.66(0.06),0.67, n=72	1.21(0.06),1.22, n=72
JUN collection	0.67(0.05),0.67, n=11	1.23(0.05),1.22, n=11
JUL collection	0.63(0.06),0.64, n=36	1.16(0.13),1.21, n=38
AUG collection	0.65(0.05),0.65, n=100	1.20(0.06),1.21, n=100
SEP collection	0.67(0.06),0.67, n=60	1.26(0.07),1.25, n=60
OCT collection	0.65(0.06),0.65, n=52	1.22(0.07),1.23, n=52
NOV collection	0.66(0.08),0.66, n=16	1.25(0.08),1.26, n=16

Table A.1 – *Continued on next page*

Table A.1 *Continued from previous page*

Categorical explanatory variable	GLU (mg/dl)	INS (μIU/ml)
Non-obese, no prior laminitis	76.1(10.5),76.1, n=339	7.1(8.6),4.8, n=339
Non-obese, prior laminitis	79.7(12.3),78.6, n=95	26.7(82.8),10.3, n=95
Obese, no prior laminitis	76.8(8.3),77.2, n=110	10.6(7.1),9.1, n=111
Obese, prior laminitis	78.0(9.4),78.0, n=64	23.4(23.6),14.5, n=64
Morgan	75.9(9.7),76.2, n=286	9.1(11.2),6.0, n=287
Arabian	82.0(12.6),80.2, n=63	10.2(10.1),6.5, n=63
Other high risk breeds	78.6(7.6),78.4, n=45	17.6(20.2),10.9, n=45
Other low risk breeds	75.7(7.4),76.8, n=16	10.8(11.4),5.8, n=16
Pony	75.6(12.6),73.8, n=96	23.1(82.4),7.9, n=96
Quarter Horse	76.0(10.5),75.8, n=57	6.4(5.3),5.1, n=57
Tennessee Walking Horse	80.2(6.8),80.0, n=45	17.8(21.4),11.9, n=45
Gelding	78.7(11.1),77.9, n=209	17.1(57.0),6.7, n=210
Mare	76.0(10.1),76.5, n=367	10.1(11.8),6.9, n=367
Stallion	77.7(7.5),74.7, n=32	10.4(20.1),5.3, n=32
Owner submitted sample	75.9(9.4),76.2, n=331	7.7(8.5),5.6, n=332
Reseacher collected sample	78.3(11.4),77.7, n=277	18.2(50.7),9.3, n=277
No L-thyroxine supplementation	76.8(10.4),76.6, n=584	12.1(35.5),6.3, n=585
Yes L-thyroxine supplementation	81.1(8.9),83.5, n=24	21.3(20.2),16.7, n=24
DEC collection	70.3(7.4),70.7, n=58	7.6(4.1),6.9, n=58
JAN collection	69.3(9.0),70.1, n=8	25.4(38.6),11.9, n=8
FEB collection	76.0(7.8),76.3, n=39	12.1(7.9),10.9, n=39
MAR collection	77.4(10.7),77.5, n=86	10.0(14.0),5.9, n=86
APR collection	78.2(8.1),77.3, n=69	8.9(13.6),5.2, n=69
MAY collection	76.7(7.8),76.6, n=72	10.0(14.5),5.7, n=72
JUN collection	78.8(6.5),78.5, n=11	12.5(12.0),9.6, n=11
JUL collection	80.4(8.4),79.9, n=38	9.8(11.6),5.3, n=38
AUG collection	79.5(14.1),79.6, n=100	16.0(64.0),5.1, n=100
SEP collection	78.7(12.1),76.8, n=60	20.7(66.2),7.7, n=60
OCT collection	75.5(9.4),75.5, n=52	14.3(11.3),11.1, n=52
NOV collection	74.6(7.6),73.4, n=15	13.3(15.5),6.5, n=16

Table A.1 – *Continued on next page*

Table A.1 *Continued from previous page*

Categorical explanatory variable	GLU OST (mg/dl)	INS OST (μIU/ml)
Non-obese, no prior laminitis	94.8(16.8),94.2, n=294	22.8(22.8),16.2, n=296
Non-obese, prior laminitis	102.0(24.7),98.2, n=69	80.2(137.5),37.3, n=68
Obese, no prior laminitis	101.2(15.8),102.2, n=98	42.7(34.7),29.7, n=98
Obese, prior laminitis	104.5(18.2),106.5, n=52	80.4(100.5),56.4, n=52
Morgan	94.9(17.2),95.4, n=268	34.3(59.9),20.3, n=268
Arabian	103.5(17.1),104.5, n=53	40.3(32.5),34.6, n=53
Other high risk breeds	97.4(16.5),91.2, n=28	49.6(50.5),29.7, n=28
Other low risk breeds	99.5(15.8),104.7, n=11	52.4(86.7),14.9, n=12
Pony	102.8(23.0),101.0, n=87	51.5(105.2),26.4, n=87
Quarter Horse	97.0(17.1),97.7, n=36	17.7(13.2),17.1, n=36
Tennessee Walking Horse	102.7(14.1),99.8, n=30	70.6(70.4),45.7, n=30
Gelding	101.0(19.2),98.3, n=159	52.5(108.4),21.1, n=159
Mare	96.8(18.3),96.8, n=324	34.4(35.5),23.2, n=324
Stallion	95.2(12.2),96.5, n=30	35.2(33.1),24.7, n=31
Owner submitted sample	97.3(17.3),97.2, n=330	29.1(28.3),19.7, n=330
Reseacher collected sample	99.1(20.1),97.4, n=183	59.6(103.6),27.1, n=184
No L-thyroxine supplementation	97.9(18.4),97.4, n=497	39.1(67.8),21.8, n=498
Yes L-thyroxine supplementation	101.3(18.4),98.7, n=16	69.5(48.3),66.1, n=16
DEC collection	89.2(12.5),89.4, n=58	28.5(24.0),22.3, n=57
JAN collection	90.2(15.4),93.0, n=6	46.1(48.9),31.4, n=6
FEB collection	97.5(16.4),97.2, n=32	39.0(29.2),34.6, n=32
MAR collection	99.8(15.8),99.8, n=78	28.4(32.5),19.6, n=78
APR collection	105.9(17.2),105.2, n=54	35.0(35.8),21.8, n=55
MAY collection	91.2(16.8),90.5, n=64	29.2(30.4),16.7, n=64
JUN collection	103.4(11.7),108.5, n=7	23.8(14.7),25.6, n=7
JUL collection	103.9(16.1),105.0, n=35	35.7(35.3),19.9, n=35
AUG collection	93.5(20.2),91.7, n=84	37.9(77.0),17.2, n=84
SEP collection	106.9(23.3),108.5, n=56	77.9(150.3),39.5, n=56
OCT collection	99.3(14.4),100.3, n=27	62.5(63.0),48.9, n=28
NOV collection	100.4(18.8),101.7, n=12	58.9(61.1),37.8, n=12

Table A.1 – *Continued on next page*

Table A.1 *Continued from previous page*

Categorical explanatory variable	TG (mg/dl)	NEFA (mmol/L)
Non-obese, no prior laminitis	24.6(14.1),21.2, n=337	0.23(0.19),0.18, n=338
Non-obese, prior laminitis	43.2(32.3),37.0, n=91	0.21(0.16),0.18, n=94
Obese, no prior laminitis	31.8(17.2),27.6, n=110	0.25(0.15),0.22, n=111
Obese, prior laminitis	53.1(53.7),38.1, n=64	0.27(0.20),0.23, n=64
Morgan	27.0(16.5),23.1, n=286	0.23(0.19),0.19, n=288
Arabian	30.3(15.8),25.7, n=61	0.21(0.15),0.20, n=62
Other high risk breeds	36.0(26.6),29.8, n=44	0.16(0.14),0.12, n=45
Other low risk breeds	23.5(11.4),21.6, n=16	0.33(0.17),0.28, n=16
Pony	47.3(51.0),31.3, n=95	0.31(0.19),0.30, n=96
Quarter Horse	27.5(16.6),22.7, n=57	0.23(0.19),0.18, n=56
Tennessee Walking Horse	35.1(18.8),33.1, n=43	0.18(0.12),0.16, n=44
Gelding	28.9(16.3),25.4, n=208	0.20(0.16),0.16, n=209
Mare	34.3(31.9),26.5, n=363	0.25(0.19),0.22, n=366
Stallion	20.4(8.4),19.2, n=31	0.24(0.20),0.22, n=32
Owner submitted sample	31.8(31.8),23.6, n=332	0.27(0.19),0.23, n=333
Reseacher collected sample	31.7(19.1),28.3, n=270	0.19(0.16),0.15, n=274
No L-thyroxine supplementation	31.0(26.3),25.0, n=578	0.24(0.18),0.20, n=583
Yes L-thyroxine supplementation	48.6(34.1),43.5, n=24	0.18(0.11),0.16, n=24
DEC collection	29.3(19.4),26.2, n=58	0.27(0.16),0.26, n=58
JAN collection	29.3(11.9),31.3, n=8	0.27(0.22),0.16, n=8
FEB collection	36.4(18.7),32.2, n=37	0.19(0.11),0.21, n=39
MAR collection	27.9(22.0),20.2, n=86	0.23(0.18),0.20, n=86
APR collection	27.6(16.5),23.1, n=68	0.28(0.19),0.23, n=69
MAY collection	28.0(15.5),24.9, n=73	0.20(0.17),0.16, n=73
JUN collection	17.9(7.4),15.0, n=11	0.08(0.08),0.06, n=11
JUL collection	33.0(17.4),25.6, n=38	0.24(0.16),0.26, n=38
AUG collection	31.2(23.7),25.0, n=97	0.22(0.22),0.16, n=97
SEP collection	52.9(59.9),33.3, n=60	0.28(0.20),0.30, n=60
OCT collection	29.3(12.0),27.6, n=50	0.21(0.16),0.15, n=52
NOV collection	23.9(11.2),26.6, n=16	0.25(0.14),0.28, n=16

Table A.1 – *Continued on next page*

Table A.1 *Continued from previous page*

Categorical explanatory variable	Leptin (ng/ml)	Adiponectin (ng/ml)
Non-obese, no prior laminitis	5.3(3.5),4.4, n=335	5017(2760),4504, n=340
Non-obese, prior laminitis	5.1(2.9),4.5, n=93	3106(3455),1613, n=95
Obese, no prior laminitis	8.3(4.5),7.8, n=109	4335(3263),3502, n=110
Obese, prior laminitis	7.5(4.1),6.9, n=63	2667(2684),1612, n=64
Morgan	6.8(4.2),5.6, n=282	4938(3018),4537, n=287
Arabian	5.5(2.8),4.9, n=63	2795(2166),2463, n=63
Other high risk breeds	7.4(4.4),7.3, n=43	3357(2514),2743, n=45
Other low risk breeds	4.2(3.4),2.9, n=16	3602(2994),3266, n=16
Pony	5.4(3.7),4.4, n=96	4944(3645),3899, n=96
Quarter Horse	3.5(2.4),2.7, n=57	4466(3042),4018, n=57
Tennessee Walking Horse	5.9(3.1),5.7, n=43	2602(2302),1750, n=45
Gelding	5.4(3.3),4.7, n=209	4064(3076),3330, n=210
Mare	6.5(4.2),5.6, n=360	4345(2984),3742, n=367
Stallion	4.4(2.8),3.5, n=31	6254(3720),5174, n=32
Owner submitted sample	6.1(4.1),5.1, n=329	5142(3050),4573, n=333
Reseacher collected sample	5.9(3.7),5.2, n=271	3392(2860),2895, n=276
No L-thyroxine supplementation	6.0(3.9),5.1, n=578	4424(3074),3772, n=585
Yes L-thyroxine supplementation	6.8(3.8),5.9, n=22	2523(2930),1333, n=24
DEC collection	6.3(4.5),5.1, n=58	5351(3191),4765, n=58
JAN collection	6.7(4.6),5.7, n=8	2171(1753),1778, n=8
FEB collection	7.3(5.1),6.1, n=38	3702(2832),3037, n=39
MAR collection	5.5(4.0),4.1, n=86	5895(3409),5638, n=86
APR collection	6.7(3.6),5.5, n=69	4671(3149),4295, n=69
MAY collection	7.5(4.0),6.8, n=73	3257(1969),3294, n=73
JUN collection	3.1(1.4),2.7, n=11	3700(3127),3333, n=11
JUL collection	5.0(2.5),4.8, n=38	3693(2414),3353, n=38
AUG collection	3.9(2.7),3.1, n=91	4449(3100),4098, n=99
SEP collection	4.9(3.1),4.0, n=60	4111(3715),3429, n=60
OCT collection	8.7(3.3),9.1, n=52	3357(2437),3050, n=52
NOV collection	6.8(2.8),7.1, n=16	4157(2441),3125, n=16

Table A.1 – *Continued on next page*

Table A.1 *Continued from previous page*

Categorical explanatory variable	ACTH (pg/ml)
Non-obese, no prior laminitis	31.0(25.0),25.1, n=339
Non-obese, prior laminitis	46.0(48.1),30.1, n=94
Obese, no prior laminitis	33.6(29.5),26.4, n=111
Obese, prior laminitis	45.8(39.2),33.3, n=64
Morgan	30.3(27.7),23.1, n=286
Arabian	45.0(44.4),32.2, n=63
Other high risk breeds	36.7(27.5),28.1, n=45
Other low risk breeds	51.5(51.6),36.8, n=16
Pony	41.8(31.7),32.5, n=96
Quarter Horse	35.9(33.1),27.0, n=57
Tennessee Walking Horse	32.6(34.0),24.5, n=45
Gelding	35.3(31.0),27.1, n=210
Mare	35.5(34.6),25.8, n=366
Stallion	34.3(12.7),33.2, n=32
Owner submitted sample	30.4(21.9),24.1, n=332
Researcher collected sample	41.3(41.2),29.4, n=276
No L-thyroxine supplementation	35.1(32.8),26.6, n=584
Yes L-thyroxine supplementation	40.9(25.8),29.0, n=24
DEC collection	26.6(12.6),22.5, n=57
JAN collection	31.6(13.3),28.0, n=8
FEB collection	26.8(12.8),22.9, n=39
MAR collection	25.3(14.2),21.6, n=86
APR collection	23.3(19.8),20.3, n=69
MAY collection	29.7(28.5),24.1, n=73
JUN collection	18.9(5.8),17.5, n=11
JUL collection	44.3(46.4),28.8, n=38
AUG collection	41.5(28.0),34.4, n=100
SEP collection	61.1(56.7),41.6, n=60
OCT collection	45.6(38.0),36.1, n=51
NOV collection	43.5(37.4),28.5, n=16

Table A.2: Exercise demographics

	N=
Type of exericse	
Dressage	43
Driving	13
Eventing/Hunter Jumper	10
Pleasure/trail	183
Longline	1
Other	7
Hours exercise/week	
0	353
0.01-3	164
3.01-5	61
> 5	32

Table A.3: Individual level covariance estimates(SE) for metabolic phenotypes from the null multivariate, multilevel model.

	NH	GH	GLU	INS	GLU OST	INS OST	TG	NEFA	ACTH	LEP	APN
NH	0.57 (0.04)										
GH	0.23 (0.03)	0.60 (0.04)									
GLU	0.00 (0.03)	0.00 (0.03)	0.69 (0.05)								
INS	0.19 (0.03)	0.14 (0.03)	0.26 (0.03)	0.69 (0.05)							
GLU OST	0.02 (0.03)	0.02 (0.03)	0.33 (0.04)	0.16 (0.03)	0.61 (0.04)						
INS OST	0.20 (0.03)	0.14 (0.03)	0.20 (0.04)	0.47 (0.04)	0.28 (0.04)	0.71 (0.05)					
TG	0.14 (0.03)	0.07 (0.03)	0.08 (0.03)	0.23 (0.03)	0.08 (0.03)	0.27 (0.04)	0.71 (0.05)				
NEFA	-0.01 (0.03)	-0.01 (0.03)	-0.09 (0.03)	-0.06 (0.03)	0.06 (0.03)	-0.04 (0.03)	-0.02 (0.03)	0.56 (0.04)			
ACTH	0.04 (0.03)	-0.01 (0.03)	0.03 (0.03)	0.09 (0.03)	0.03 (0.03)	0.05 (0.03)	0.04 (0.03)	0.05 (0.03)	0.56 (0.04)		
LEP	0.18 (0.03)	0.28 (0.03)	0.11 (0.03)	0.24 (0.03)	0.13 (0.03)	0.29 (0.04)	0.16 (0.03)	-0.11 (0.03)	-0.01 (0.03)	0.71 (0.05)	
APN	-0.14 (0.03)	-0.06 (0.03)	-0.07 (0.03)	-0.26 (0.04)	-0.05 (0.03)	-0.27 (0.04)	-0.31 (0.04)	0.08 (0.03)	0.00 (0.03)	-0.09 (0.03)	0.78 (0.05)

All phenotypes were scaled to zero mean and variance equal to one standard deviation. INS, INS OST, TG, and ACTH were log transformed. NEFA, LEP, and APN were square root transformed.

Table A.4: Farm level covariance estimates(SE) for metabolic phenotypes from the null multivariate, multilevel model.

	NH	GH	GLU	INS	GLU OST	INS OST	TG	NEFA	ACTH	LEP	APN
NH	0.50 (0.08)										
GH	0.24 (0.06)	0.45 (0.08)									
GLU	-0.01 (0.06)	-0.04 (0.05)	0.37 (0.07)								
INS	0.00 (0.06)	0.00 (0.05)	0.12 (0.06)	0.39 (0.08)							
GLU OST	0.05 (0.06)	0.06 (0.06)	0.32 (0.07)	0.18 (0.06)	0.45 (0.09)						
INS OST	-0.01 (0.06)	0.03 (0.06)	0.14 (0.06)	0.32 (0.07)	0.28 (0.07)	0.35 (0.07)					
TG	0.02 (0.05)	0.01 (0.05)	0.04 (0.04)	0.13 (0.05)	0.08 (0.05)	0.08 (0.05)	0.22 (0.05)				
NEFA	0.04 (0.06)	0.12 (0.06)	0.01 (0.05)	-0.03 (0.05)	0.17 (0.06)	0.09 (0.05)	0.03 (0.04)	0.45 (0.08)			
ACTH	-0.09 (0.06)	-0.06 (0.06)	0.12 (0.06)	0.16 (0.06)	0.17 (0.06)	0.23 (0.06)	0.11 (0.05)	0.13 (0.06)	0.54 (0.09)		
LEP	0.05 (0.05)	0.03 (0.05)	-0.08 (0.05)	0.17 (0.05)	0.02 (0.05)	0.15 (0.05)	0.04 (0.04)	0.04 (0.05)	0.03 (0.06)	0.33 (0.07)	
APN	0.13 (0.05)	0.09 (0.05)	-0.16 (0.05)	-0.16 (0.05)	-0.14 (0.05)	-0.14 (0.05)	-0.11 (0.04)	-0.01 (0.05)	-0.16 (0.05)	-0.03 (0.04)	0.23 (0.05)

Table A.5: Fixed effect estimates for neck circumference to height ratio (NH) from the full multivariate, multilevel model.

Parameter	scaled estimate(SE)	unscaled estimate(SE)	p-value
Intercept	0.235(0.296)	0.670(0.016)	
+1 SD age(yrs)	0.013(0.033)	0.001(0.002)	7.0e-01
Sex:gelding	Reference	Reference	Reference
Sex:mare	-0.197(0.073)	-0.011(0.004)	**7.1e-03**
Sex:stallion	0.518(0.149)	0.028(0.008)	**5.3e-04**
Breed:Morgan	Reference	Reference	Reference
Breed:Arab	-0.662(0.186)	-0.035(0.010)	**3.8e-04**
Breed:Pony	-0.002(0.189)	0.000(0.010)	9.9e-01
Breed:TW	-0.214(0.213)	-0.011(0.011)	3.1e-01
Breed:QH	-0.175(0.182)	-0.009(0.010)	3.4e-01
Breed:HR	-0.484(0.205)	-0.026(0.011)	**1.8e-02**
Breed:LR	-0.855(0.256)	-0.046(0.014)	**8.4e-04**
Group:obese$^-$laminitis$^-$	Reference	Reference	Reference
Group:obese$^-$laminitis$^+$	0.331(0.092)	0.018(0.005)	**3.2e-04**
Group:obese$^+$laminitis$^-$	0.652(0.086)	0.035(0.005)	**3.2e-14**
Group:obese$^+$laminitis$^+$	0.957(0.121)	0.051(0.006)	**2.6e-15**
+1 SD g Mcal/kg bwt per day	0.192(0.340)	0.010(0.018)	5.7e-01
+1 SD g CP/kg bwt per day	-0.095(0.161)	-0.005(0.009)	5.5e-01
+1 SD g NDF/kg bwt per day	-0.115(0.192)	-0.006(0.010)	5.5e-01
+1 SD g Starch/kg bwt per day	-0.065(0.081)	-0.003(0.004)	4.2e-01
+1 SD g WSC/kg bwt per day	-0.089(0.109)	-0.005(0.006)	4.1e-01
+1 SD hours grazing per day	0.121(0.063)	0.006(0.003)	5.3e-02
+1 SD hours exercise per day	0.067(0.041)	0.004(0.002)	1.1e-01
+1 SD hours stalled per day	-0.046(0.043)	-0.002(0.002)	2.8e-01
Month:DEC	Reference	Reference	Reference
Month:JAN	-0.197(0.529)	-0.011(0.028)	7.1e-01
Month:FEB	0.353(0.383)	0.019(0.020)	3.6e-01
Month:MAR	-0.492(0.356)	-0.026(0.019)	1.7e-01
Month:APR	-0.373(0.379)	-0.020(0.020)	3.2e-01
Month:MAY	0.055(0.358)	0.003(0.019)	8.8e-01
Month:JUN	0.490(0.505)	0.026(0.027)	3.3e-01
Month:JUL	-0.315(0.425)	-0.017(0.023)	4.6e-01
Month:AUG	-0.227(0.339)	-0.012(0.018)	5.0e-01
Month:SEP	-0.081(0.358)	-0.004(0.019)	8.2e-01
Month:OCT	-0.241(0.351)	-0.013(0.019)	4.9e-01
Month:NOV	-0.042(0.456)	-0.002(0.024)	9.3e-01
+1 SD latitude	0.019(0.079)	0.001(0.004)	8.1e-01
Oral thyroxine (no)	Reference	Reference	Reference
Oral thyroxine (yes)	0.025(0.178)	0.001(0.009)	8.9e-01
Researcher collected sample	Reference	Reference	Reference
Owner submitted sample	-0.368(0.204)	-0.020(0.011)	7.1e-02

Metabolic traits were mean centered and variance scaled to 1 standard deviation. Scaled fixed effect estimates represent the number of standard deviations change in the metabolic trait relative to the reference group.

Table A.6: Fixed effect estimates for girth to height ratio (GH) from the full multivariate, multilevel model.

Parameter	scaled estimate(SE)	unscaled estimate(SE)	p-value
Intercept	0.026(0.250)	1.223(0.018)	
+1 SD age(yrs)	0.065(0.033)	0.005(0.002)	5.2e-02
Sex:gelding	Reference	Reference	Reference
Sex:mare	0.134(0.073)	0.010(0.005)	**6.7e-02**
Sex:stallion	-0.438(0.153)	-0.031(0.011)	**4.1e-03**
Breed:Morgan	Reference	Reference	Reference
Breed:Arab	-0.213(0.175)	-0.015(0.012)	2.2e-01
Breed:Pony	-0.055(0.172)	-0.004(0.012)	7.5e-01
Breed:TW	-0.246(0.196)	-0.017(0.014)	2.1e-01
Breed:QH	0.301(0.168)	0.021(0.012)	7.3e-02
Breed:HR	-0.268(0.190)	-0.019(0.013)	1.6e-01
Breed:LR	-0.468(0.245)	-0.033(0.017)	5.6e-02
Group:obese$^-$laminitis$^-$	Reference	Reference	Reference
Group:obese$^-$laminitis$^+$	-0.026(0.093)	-0.002(0.007)	7.8e-01
Group:obese$^+$laminitis$^-$	0.666(0.087)	0.047(0.006)	**1.5e-14**
Group:obese$^+$laminitis$^+$	0.651(0.120)	0.046(0.009)	**6.2e-08**
+1 SD g Mcal/kg bwt per day	-0.530(0.317)	-0.038(0.022)	9.5e-02
+1 SD g CP/kg bwt per day	0.199(0.147)	0.014(0.010)	1.8e-01
+1 SD g NDF/kg bwt per day	0.298(0.179)	0.021(0.013)	9.5e-02
+1 SD g Starch/kg bwt per day	0.097(0.077)	0.007(0.005)	2.1e-01
+1 SD g WSC/kg bwt per day	-0.120(0.101)	-0.008(0.007)	2.4e-01
+1 SD hours grazing per day	0.095(0.059)	0.007(0.004)	1.1e-01
+1 SD hours exercise per day	-0.013(0.040)	-0.001(0.003)	7.5e-01
+1 SD hours stalled per day	-0.045(0.041)	-0.003(0.003)	2.7e-01
Month:DEC	Reference	Reference	Reference
Month:JAN	0.330(0.467)	0.023(0.033)	4.8e-01
Month:FEB	0.082(0.322)	0.006(0.023)	8.0e-01
Month:MAR	-0.244(0.301)	-0.017(0.021)	4.2e-01
Month:APR	-0.306(0.321)	-0.022(0.023)	3.4e-01
Month:MAY	-0.024(0.305)	-0.002(0.022)	9.4e-01
Month:JUN	0.220(0.437)	0.016(0.031)	6.1e-01
Month:JUL	-1.447(0.358)	-0.103(0.025)	**5.3e-05**
Month:AUG	-0.238(0.288)	-0.017(0.020)	4.1e-01
Month:SEP	0.082(0.305)	0.006(0.022)	7.9e-01
Month:OCT	-0.170(0.300)	-0.012(0.021)	5.7e-01
Month:NOV	0.309(0.391)	0.022(0.028)	4.3e-01
+1 SD latitude	0.100(0.069)	0.007(0.005)	1.5e-01
Oral thyroxine (no)	Reference	Reference	Reference
Oral thyroxine (yes)	-0.143(0.176)	-0.010(0.012)	4.2e-01
Researcher collected sample	Reference	Reference	Reference
Owner submitted sample	-0.183(0.175)	-0.013(0.012)	2.9e-01

Metabolic traits were mean centered and variance scaled to 1 standard deviation. Scaled fixed effect estimates represent the number of standard deviations change in the metabolic trait relative to the reference group.

Table A.7: Fixed effect estimates for fasting glucose (mg/dl) from the full multivariate, multilevel model.

Parameter	scaled estimate(SE)	unscaled estimate(SE)	p-value
Intercept	-0.711(0.258)	69.581(2.685)	
+1 SD age(yrs)	-0.036(0.040)	-0.380(0.413)	3.6e-01
Sex:gelding	Reference	Reference	Reference
Sex:mare	-0.161(0.087)	-1.680(0.905)	6.3e-02
Sex:stallion	0.365(0.184)	3.806(1.914)	**4.7e-02**
Breed:Morgan	Reference	Reference	Reference
Breed:Arab	0.387(0.197)	4.027(2.052)	5.0e-02
Breed:Pony	0.117(0.189)	1.215(1.973)	5.4e-01
Breed:TW	0.148(0.217)	1.546(2.264)	4.9e-01
Breed:QH	-0.158(0.189)	-1.642(1.973)	4.1e-01
Breed:HR	0.240(0.213)	2.499(2.217)	2.6e-01
Breed:LR	-0.092(0.282)	-0.955(2.940)	7.5e-01
Group:obese$^-$laminitis$^-$	Reference	Reference	Reference
Group:obese$^-$laminitis$^+$	0.153(0.111)	1.590(1.157)	1.7e-01
Group:obese$^+$laminitis$^-$	0.046(0.103)	0.478(1.076)	6.6e-01
Group:obese$^+$laminitis$^+$	0.087(0.141)	0.905(1.472)	5.4e-01
+1 SD g Mcal/kg bwt per day	-0.086(0.356)	-0.899(3.713)	8.1e-01
+1 SD g CP/kg bwt per day	0.003(0.162)	0.035(1.691)	9.8e-01
+1 SD g NDF/kg bwt per day	-0.088(0.202)	-0.913(2.099)	6.6e-01
+1 SD g Starch/kg bwt per day	-0.113(0.088)	-1.176(0.918)	2.0e-01
+1 SD g WSC/kg bwt per day	0.161(0.113)	1.682(1.180)	1.5e-01
+1 SD hours grazing per day	0.072(0.067)	0.750(0.702)	2.9e-01
+1 SD hours exercise per day	-0.043(0.046)	-0.448(0.483)	3.5e-01
+1 SD hours stalled per day	-0.021(0.048)	-0.220(0.495)	6.6e-01
Month:DEC	Reference	Reference	Reference
Month:JAN	-0.365(0.484)	-3.803(5.045)	4.5e-01
Month:FEB	0.452(0.334)	4.710(3.475)	1.8e-01
Month:MAR	0.493(0.311)	5.140(3.239)	1.1e-01
Month:APR	0.606(0.332)	6.317(3.456)	6.8e-02
Month:MAY	0.700(0.315)	7.295(3.279)	**2.6e-02**
Month:JUN	0.475(0.461)	4.944(4.800)	3.0e-01
Month:JUL	0.405(0.375)	4.218(3.905)	2.8e-01
Month:AUG	1.144(0.298)	11.918(3.105)	**1.2e-04**
Month:SEP	0.740(0.318)	7.710(3.308)	**2.0e-02**
Month:OCT	0.331(0.314)	3.444(3.267)	2.9e-01
Month:NOV	0.033(0.410)	0.340(4.274)	9.4e-01
+1 SD latitude	-0.033(0.074)	-0.348(0.771)	6.5e-01
Oral thyroxine (no)	Reference	Reference	Reference
Oral thyroxine (yes)	0.244(0.206)	2.542(2.149)	2.4e-01
Researcher collected sample	Reference	Reference	Reference
Owner submitted sample	0.170(0.184)	1.768(1.915)	3.6e-01

Metabolic traits were mean centered and variance scaled to 1 standard deviation. Scaled fixed effect estimates represent the number of standard deviations change in the metabolic trait relative to the reference group.

Table A.8: Fixed effect estimates for fasting log insulin (μIU/ml) from the full multivariate, multilevel model.

Parameter	scaled estimate(SE)	unscaled estimate(SE)	p-value
Intercept	-0.660(0.246)	0.568(0.101)	
+1 SD age(yrs)	0.121(0.036)	0.050(0.015)	**7.4e-04**
Sex:gelding	Reference	Reference	Reference
Sex:mare	0.042(0.078)	0.017(0.032)	5.9e-01
Sex:stallion	0.285(0.164)	0.117(0.067)	8.2e-02
Breed:Morgan	Reference	Reference	Reference
Breed:Arab	-0.432(0.183)	-0.178(0.075)	**1.8e-02**
Breed:Pony	0.340(0.178)	0.140(0.073)	5.6e-02
Breed:TW	0.066(0.202)	0.027(0.083)	7.4e-01
Breed:QH	-0.681(0.175)	-0.280(0.072)	**1.0e-04**
Breed:HR	0.071(0.197)	0.029(0.081)	7.2e-01
Breed:LR	-0.247(0.259)	-0.102(0.106)	3.4e-01
Group:obese$^-$laminitis$^-$	Reference	Reference	Reference
Group:obese$^-$laminitis$^+$	0.590(0.100)	0.242(0.041)	**3.5e-09**
Group:obese$^+$laminitis$^-$	0.454(0.093)	0.187(0.038)	**1.0e-06**
Group:obese$^+$laminitis$^+$	0.961(0.128)	0.395(0.053)	**5.9e-14**
+1 SD g Mcal/kg bwt per day	-0.007(0.332)	-0.003(0.136)	9.8e-01
+1 SD g CP/kg bwt per day	-0.048(0.152)	-0.020(0.063)	7.5e-01
+1 SD g NDF/kg bwt per day	-0.223(0.187)	-0.091(0.077)	2.3e-01
+1 SD g Starch/kg bwt per day	0.026(0.081)	0.011(0.033)	7.5e-01
+1 SD g WSC/kg bwt per day	0.180(0.106)	0.074(0.043)	8.8e-02
+1 SD hours grazing per day	0.056(0.062)	0.023(0.026)	3.7e-01
+1 SD hours exercise per day	-0.018(0.042)	-0.007(0.017)	6.7e-01
+1 SD hours stalled per day	-0.030(0.044)	-0.012(0.018)	4.9e-01
Month:DEC	Reference	Reference	Reference
Month:JAN	0.128(0.455)	0.052(0.187)	7.8e-01
Month:FEB	0.497(0.317)	0.204(0.130)	1.2e-01
Month:MAR	0.197(0.297)	0.081(0.122)	5.1e-01
Month:APR	-0.161(0.316)	-0.066(0.130)	6.1e-01
Month:MAY	0.202(0.299)	0.083(0.123)	5.0e-01
Month:JUN	-0.302(0.434)	-0.124(0.178)	4.9e-01
Month:JUL	-0.094(0.355)	-0.038(0.146)	7.9e-01
Month:AUG	0.104(0.284)	0.043(0.117)	7.1e-01
Month:SEP	-0.050(0.302)	-0.021(0.124)	8.7e-01
Month:OCT	0.173(0.297)	0.071(0.122)	5.6e-01
Month:NOV	-0.182(0.387)	-0.075(0.159)	6.4e-01
+1 SD latitude	0.066(0.069)	0.027(0.029)	3.4e-01
Oral thyroxine (no)	Reference	Reference	Reference
Oral thyroxine (yes)	0.107(0.187)	0.044(0.077)	5.7e-01
Researcher collected sample	Reference	Reference	Reference
Owner submitted sample	0.667(0.174)	0.274(0.071)	**1.2e-04**

Metabolic traits were mean centered and variance scaled to 1 standard deviation. Scaled fixed effect estimates represent the number of standard deviations change in the metabolic trait relative to the reference group.

Table A.9: Fixed effect estimates for 75 minute post oral sugar test glucose (mg/dl) from the full multivariate, multilevel model.

Parameter	scaled estimate(SE)	unscaled estimate(SE)	p-value
Intercept	-0.662(0.281)	85.827(5.159)	
+1 SD age(yrs)	-0.020(0.040)	-0.374(0.729)	6.1e-01
Sex:gelding	Reference	Reference	Reference
Sex:mare	-0.103(0.089)	-1.884(1.636)	2.5e-01
Sex:stallion	0.370(0.181)	6.803(3.317)	**4.0e-02**
Breed:Morgan	Reference	Reference	Reference
Breed:Arab	0.064(0.211)	1.184(3.884)	7.6e-01
Breed:Pony	0.379(0.206)	6.958(3.779)	6.6e-02
Breed:TW	0.158(0.250)	2.894(4.586)	5.3e-01
Breed:QH	-0.086(0.219)	-1.584(4.017)	6.9e-01
Breed:HR	-0.012(0.250)	-0.212(4.588)	9.6e-01
Breed:LR	0.089(0.309)	1.636(5.683)	7.7e-01
Group:obese$^-$laminitis$^-$	Reference	Reference	Reference
Group:obese$^-$laminitis$^+$	0.096(0.116)	1.763(2.138)	4.1e-01
Group:obese$^+$laminitis$^-$	0.164(0.103)	3.006(1.883)	1.1e-01
Group:obese$^+$laminitis$^+$	0.184(0.148)	3.386(2.720)	2.1e-01
+1 SD g Mcal/kg bwt per day	0.262(0.381)	4.817(6.992)	4.9e-01
+1 SD g CP/kg bwt per day	-0.181(0.173)	-3.329(3.187)	3.0e-01
+1 SD g NDF/kg bwt per day	-0.259(0.222)	-4.752(4.078)	2.4e-01
+1 SD g Starch/kg bwt per day	-0.190(0.091)	-3.489(1.674)	**3.7e-02**
+1 SD g WSC/kg bwt per day	0.120(0.123)	2.208(2.252)	3.3e-01
+1 SD hours grazing per day	0.008(0.072)	0.148(1.331)	9.1e-01
+1 SD hours exercise per day	0.018(0.053)	0.332(0.979)	7.3e-01
+1 SD hours stalled per day	-0.125(0.049)	-2.304(0.904)	**1.1e-02**
Month:DEC	Reference	Reference	Reference
Month:JAN	-0.217(0.545)	-3.988(10.016)	6.9e-01
Month:FEB	0.559(0.369)	10.266(6.775)	1.3e-01
Month:MAR	0.327(0.341)	5.997(6.256)	3.4e-01
Month:APR	0.540(0.379)	9.914(6.960)	1.5e-01
Month:MAY	0.573(0.348)	10.524(6.398)	1.0e-01
Month:JUN	0.639(0.539)	11.738(9.897)	2.4e-01
Month:JUL	0.264(0.411)	4.853(7.541)	5.2e-01
Month:AUG	0.600(0.329)	11.028(6.042)	6.8e-02
Month:SEP	0.770(0.347)	14.133(6.366)	**2.6e-02**
Month:OCT	0.516(0.358)	9.481(6.573)	1.5e-01
Month:NOV	0.279(0.464)	5.122(8.526)	5.5e-01
+1 SD latitude	-0.017(0.086)	-0.305(1.579)	8.5e-01
Oral thyroxine (no)	Reference	Reference	Reference
Oral thyroxine (yes)	0.207(0.229)	3.798(4.205)	3.7e-01
Researcher collected sample	Reference	Reference	Reference
Owner submitted sample	0.092(0.206)	1.697(3.777)	6.5e-01

Metabolic traits were mean centered and variance scaled to 1 standard deviation. Scaled fixed effect estimates represent the number of standard deviations change in the metabolic trait relative to the reference group.

Table A.10: Fixed effect estimates for 75 minute post oral sugar test log insulin (μIU/ml) from the full multivariate, multilevel model.

Parameter	scaled estimate(SE)	unscaled estimate(SE)	p-value
Intercept	-0.399(0.253)	1.173(0.115)	
+1 SD age(yrs)	0.095(0.037)	0.043(0.017)	**1.1e-02**
Sex:gelding	Reference	Reference	Reference
Sex:mare	0.129(0.083)	0.059(0.038)	1.2e-01
Sex:stallion	0.667(0.169)	0.303(0.077)	**7.5e-05**
Breed:Morgan	Reference	Reference	Reference
Breed:Arab	-0.231(0.195)	-0.105(0.089)	2.4e-01
Breed:Pony	0.107(0.189)	0.049(0.086)	5.7e-01
Breed:TW	0.299(0.229)	0.136(0.104)	1.9e-01
Breed:QH	-0.769(0.201)	-0.349(0.091)	**1.3e-04**
Breed:HR	0.104(0.229)	0.047(0.104)	6.5e-01
Breed:LR	-0.237(0.285)	-0.108(0.129)	4.1e-01
Group:obese$^-$laminitis$^-$	Reference	Reference	Reference
Group:obese$^-$laminitis$^+$	0.546(0.109)	0.248(0.049)	**5.4e-07**
Group:obese$^+$laminitis$^-$	0.483(0.096)	0.219(0.044)	**5.0e-07**
Group:obese$^+$laminitis$^+$	0.741(0.138)	0.337(0.062)	**7.2e-08**
+1 SD g Mcal/kg bwt per day	-0.237(0.351)	-0.108(0.159)	5.0e-01
+1 SD g CP/kg bwt per day	0.087(0.159)	0.040(0.072)	5.8e-01
+1 SD g NDF/kg bwt per day	-0.079(0.204)	-0.036(0.093)	7.0e-01
+1 SD g Starch/kg bwt per day	-0.062(0.084)	-0.028(0.038)	4.6e-01
+1 SD g WSC/kg bwt per day	0.134(0.113)	0.061(0.051)	2.4e-01
+1 SD hours grazing per day	-0.007(0.067)	-0.003(0.030)	9.2e-01
+1 SD hours exercise per day	0.011(0.049)	0.005(0.022)	8.2e-01
+1 SD hours stalled per day	-0.021(0.046)	-0.010(0.021)	6.5e-01
Month:DEC	Reference	Reference	Reference
Month:JAN	-0.420(0.495)	-0.191(0.225)	4.0e-01
Month:FEB	0.256(0.333)	0.116(0.151)	4.4e-01
Month:MAR	-0.282(0.307)	-0.128(0.139)	3.6e-01
Month:APR	-0.145(0.342)	-0.066(0.155)	6.7e-01
Month:MAY	-0.096(0.314)	-0.044(0.143)	7.6e-01
Month:JUN	-0.728(0.489)	-0.331(0.222)	1.4e-01
Month:JUL	-0.415(0.371)	-0.189(0.169)	2.6e-01
Month:AUG	-0.249(0.296)	-0.113(0.135)	4.0e-01
Month:SEP	-0.032(0.313)	-0.015(0.142)	9.2e-01
Month:OCT	0.175(0.323)	0.080(0.147)	5.9e-01
Month:NOV	-0.172(0.420)	-0.078(0.191)	6.8e-01
+1 SD latitude	0.030(0.078)	0.013(0.036)	7.1e-01
Oral thyroxine (no)	Reference	Reference	Reference
Oral thyroxine (yes)	0.122(0.212)	0.056(0.096)	5.6e-01
Researcher collected sample	Reference	Reference	Reference
Owner submitted sample	0.497(0.186)	0.226(0.085)	**7.5e-03**

Metabolic traits were mean centered and variance scaled to 1 standard deviation. Scaled fixed effect estimates represent the number of standard deviations change in the metabolic trait relative to the reference group.

Table A.11: Fixed effect estimates for fasting log triglycerides (mg/dl) from the full multivariate, multilevel model.

Parameter	scaled estimate(SE)	unscaled estimate(SE)	p-value
Intercept	-0.223(0.209)	1.359(0.054)	
+1 SD age(yrs)	-0.070(0.037)	-0.018(0.009)	5.4e-02
Sex:gelding	Reference	Reference	Reference
Sex:mare	0.188(0.080)	0.048(0.021)	**1.9e-02**
Sex:stallion	0.164(0.173)	0.042(0.045)	3.4e-01
Breed:Morgan	Reference	Reference	Reference
Breed:Arab	-0.129(0.175)	-0.033(0.045)	4.6e-01
Breed:Pony	0.104(0.163)	0.027(0.042)	5.2e-01
Breed:TW	0.230(0.194)	0.059(0.050)	2.4e-01
Breed:QH	-0.307(0.167)	-0.079(0.043)	6.6e-02
Breed:HR	0.254(0.189)	0.066(0.049)	1.8e-01
Breed:LR	-0.414(0.253)	-0.107(0.065)	1.0e-01
Group:obese$^-$laminitis$^-$	Reference	Reference	Reference
Group:obese$^-$laminitis$^+$	0.576(0.105)	0.149(0.027)	**3.8e-08**
Group:obese$^+$laminitis$^-$	0.205(0.096)	0.053(0.025)	**3.2e-02**
Group:obese$^+$laminitis$^+$	0.599(0.129)	0.155(0.033)	**3.6e-06**
+1 SD g Mcal/kg bwt per day	-0.427(0.309)	-0.110(0.080)	1.7e-01
+1 SD g CP/kg bwt per day	0.119(0.138)	0.031(0.036)	3.9e-01
+1 SD g NDF/kg bwt per day	0.131(0.175)	0.034(0.045)	4.5e-01
+1 SD g Starch/kg bwt per day	-0.040(0.078)	-0.010(0.020)	6.0e-01
+1 SD g WSC/kg bwt per day	0.065(0.098)	0.017(0.025)	5.0e-01
+1 SD hours grazing per day	0.066(0.060)	0.017(0.016)	2.7e-01
+1 SD hours exercise per day	0.006(0.042)	0.002(0.011)	8.8e-01
+1 SD hours stalled per day	-0.018(0.043)	-0.005(0.011)	6.7e-01
Month:DEC	Reference	Reference	Reference
Month:JAN	-0.332(0.407)	-0.086(0.105)	4.1e-01
Month:FEB	0.298(0.276)	0.077(0.071)	2.8e-01
Month:MAR	-0.266(0.252)	-0.069(0.065)	2.9e-01
Month:APR	-0.273(0.269)	-0.071(0.070)	3.1e-01
Month:MAY	-0.062(0.257)	-0.016(0.066)	8.1e-01
Month:JUN	-1.301(0.386)	-0.336(0.100)	**7.5e-04**
Month:JUL	-0.044(0.309)	-0.011(0.080)	8.9e-01
Month:AUG	-0.024(0.243)	-0.006(0.063)	9.2e-01
Month:SEP	-0.047(0.260)	-0.012(0.067)	8.6e-01
Month:OCT	-0.325(0.260)	-0.084(0.067)	2.1e-01
Month:NOV	-0.789(0.338)	-0.204(0.087)	**2.0e-02**
+1 SD latitude	-0.050(0.063)	-0.013(0.016)	4.3e-01
Oral thyroxine (no)	Reference	Reference	Reference
Oral thyroxine (yes)	0.437(0.189)	0.113(0.049)	**2.1e-02**
Researcher collected sample	Reference	Reference	Reference
Owner submitted sample	0.127(0.153)	0.033(0.040)	4.1e-01

Metabolic traits were mean centered and variance scaled to 1 standard deviation. Scaled fixed effect estimates represent the number of standard deviations change in the metabolic trait relative to the reference group.

Table A.12: Fixed effect estimates for fasting $\surd$NEFA (mmol/L) from the full multivariate, multilevel model.

Parameter	scaled estimate(SE)	unscaled estimate(SE)	p-value
Intercept	0.235(0.268)	0.490(0.051)	
+1 SD age(yrs)	-0.004(0.036)	-0.001(0.007)	9.2e-01
Sex:gelding	Reference	Reference	Reference
Sex:mare	0.071(0.080)	0.013(0.015)	3.8e-01
Sex:stallion	0.159(0.167)	0.030(0.032)	3.4e-01
Breed:Morgan	Reference	Reference	Reference
Breed:Arab	0.051(0.192)	0.010(0.037)	7.9e-01
Breed:Pony	0.201(0.189)	0.038(0.036)	2.9e-01
Breed:TW	0.138(0.215)	0.026(0.041)	5.2e-01
Breed:QH	0.232(0.185)	0.044(0.035)	2.1e-01
Breed:HR	0.142(0.207)	0.027(0.039)	4.9e-01
Breed:LR	0.579(0.267)	0.110(0.051)	**3.0e-02**
Group:obese$^-$laminitis$^-$	Reference	Reference	Reference
Group:obese$^-$laminitis$^+$	0.043(0.102)	0.008(0.019)	6.7e-01
Group:obese$^+$laminitis$^-$	0.000(0.095)	0.000(0.018)	1.0e+00
Group:obese$^+$laminitis$^+$	0.088(0.131)	0.017(0.025)	5.0e-01
+1 SD g Mcal/kg bwt per day	0.009(0.347)	0.002(0.066)	9.8e-01
+1 SD g CP/kg bwt per day	-0.141(0.160)	-0.027(0.030)	3.8e-01
+1 SD g NDF/kg bwt per day	0.101(0.195)	0.019(0.037)	6.1e-01
+1 SD g Starch/kg bwt per day	0.074(0.085)	0.014(0.016)	3.8e-01
+1 SD g WSC/kg bwt per day	-0.012(0.111)	-0.002(0.021)	9.1e-01
+1 SD hours grazing per day	0.286(0.066)	0.054(0.012)	**1.3e-05**
+1 SD hours exercise per day	0.003(0.044)	0.001(0.008)	9.4e-01
+1 SD hours stalled per day	-0.122(0.045)	-0.023(0.009)	**6.7e-03**
Month:DEC	Reference	Reference	Reference
Month:JAN	-0.090(0.492)	-0.017(0.093)	8.5e-01
Month:FEB	-0.426(0.345)	-0.081(0.066)	2.2e-01
Month:MAR	-0.491(0.324)	-0.093(0.061)	1.3e-01
Month:APR	-0.285(0.345)	-0.054(0.065)	4.1e-01
Month:MAY	-0.597(0.326)	-0.113(0.062)	**6.7e-02**
Month:JUN	-1.372(0.470)	-0.260(0.089)	**3.5e-03**
Month:JUL	-0.821(0.386)	-0.156(0.073)	**3.3e-02**
Month:AUG	-0.305(0.310)	-0.058(0.059)	3.2e-01
Month:SEP	-0.594(0.328)	-0.113(0.062)	7.0e-02
Month:OCT	-0.504(0.323)	-0.096(0.061)	1.2e-01
Month:NOV	-0.152(0.420)	-0.029(0.080)	7.2e-01
+1 SD latitude	0.105(0.074)	0.020(0.014)	1.6e-01
Oral thyroxine (no)	Reference	Reference	Reference
Oral thyroxine (yes)	0.154(0.192)	0.029(0.036)	4.2e-01
Researcher collected sample	Reference	Reference	Reference
Owner submitted sample	-0.112(0.188)	-0.021(0.036)	5.5e-01

Metabolic traits were mean centered and variance scaled to 1 standard deviation. Scaled fixed effect estimates represent the number of standard deviations change in the metabolic trait relative to the reference group.

Table A.13: Fixed effect estimates for fasting log ACTH (pg/ml) from the full multivariate, multilevel model.

Parameter	scaled estimate(SE)	unscaled estimate(SE)	p-value
Intercept	-0.691(0.221)	1.303(0.052)	
+1 SD age(yrs)	0.302(0.033)	0.071(0.008)	**9.6e-20**
Sex:gelding	Reference	Reference	Reference
Sex:mare	-0.015(0.073)	-0.004(0.017)	8.3e-01
Sex:stallion	0.190(0.154)	0.045(0.036)	2.2e-01
Breed:Morgan	Reference	Reference	Reference
Breed:Arab	0.233(0.166)	0.055(0.039)	1.6e-01
Breed:Pony	0.588(0.159)	0.138(0.037)	**2.2e-04**
Breed:TW	-0.184(0.184)	-0.043(0.043)	3.2e-01
Breed:QH	-0.211(0.160)	-0.050(0.038)	1.9e-01
Breed:HR	0.314(0.180)	0.074(0.042)	8.1e-02
Breed:LR	0.075(0.237)	0.018(0.056)	7.5e-01
Group:obese$^-$laminitis$^-$	Reference	Reference	Reference
Group:obese$^-$laminitis$^+$	0.133(0.094)	0.031(0.022)	1.6e-01
Group:obese$^+$laminitis$^-$	0.088(0.087)	0.021(0.020)	3.1e-01
Group:obese$^+$laminitis$^+$	0.015(0.119)	0.004(0.028)	9.0e-01
+1 SD g Mcal/kg bwt per day	-0.044(0.298)	-0.010(0.070)	8.8e-01
+1 SD g CP/kg bwt per day	0.091(0.136)	0.021(0.032)	5.0e-01
+1 SD g NDF/kg bwt per day	0.040(0.169)	0.009(0.040)	8.1e-01
+1 SD g Starch/kg bwt per day	0.024(0.074)	0.006(0.017)	7.4e-01
+1 SD g WSC/kg bwt per day	0.013(0.095)	0.003(0.022)	8.9e-01
+1 SD hours grazing per day	-0.005(0.057)	-0.001(0.013)	9.4e-01
+1 SD hours exercise per day	-0.032(0.039)	-0.008(0.009)	4.2e-01
+1 SD hours stalled per day	-0.055(0.040)	-0.013(0.009)	1.7e-01
Month:DEC	Reference	Reference	Reference
Month:JAN	0.104(0.414)	0.024(0.098)	8.0e-01
Month:FEB	-0.064(0.285)	-0.015(0.067)	8.2e-01
Month:MAR	-0.076(0.266)	-0.018(0.063)	7.7e-01
Month:APR	-0.168(0.284)	-0.039(0.067)	5.5e-01
Month:MAY	0.187(0.269)	0.044(0.063)	4.9e-01
Month:JUN	-0.851(0.394)	-0.200(0.093)	**3.1e-02**
Month:JUL	0.829(0.320)	0.195(0.075)	**9.6e-03**
Month:AUG	0.863(0.255)	0.203(0.060)	**7.1e-04**
Month:SEP	1.050(0.272)	0.247(0.064)	**1.1e-04**
Month:OCT	0.793(0.268)	0.187(0.063)	**3.1e-03**
Month:NOV	0.618(0.350)	0.146(0.082)	7.7e-02
+1 SD latitude	0.141(0.063)	0.033(0.015)	**2.6e-02**
Oral thyroxine (no)	Reference	Reference	Reference
Oral thyroxine (yes)	-0.161(0.173)	-0.038(0.041)	3.5e-01
Researcher collected sample	Reference	Reference	Reference
Owner submitted sample	0.437(0.157)	0.103(0.037)	**5.4e-03**

Metabolic traits were mean centered and variance scaled to 1 standard deviation. Scaled fixed effect estimates represent the number of standard deviations change in the metabolic trait relative to the reference group.

Table A.14: Fixed effect estimates for fasting $\surd$leptin (ng/ml) from the full multivariate, multilevel model.

Parameter	scaled estimate(SE)	unscaled estimate(SE)	p-value
Intercept	-0.041(0.198)	2.296(0.154)	
+1 SD age(yrs)	0.000(0.036)	0.000(0.028)	9.9e-01
Sex:gelding	Reference	Reference	Reference
Sex:mare	0.220(0.079)	0.171(0.061)	**5.2e-03**
Sex:stallion	-0.114(0.170)	-0.089(0.132)	5.0e-01
Breed:Morgan	Reference	Reference	Reference
Breed:Arab	-0.462(0.167)	-0.359(0.130)	**5.8e-03**
Breed:Pony	-0.307(0.154)	-0.239(0.120)	**4.6e-02**
Breed:TW	-0.054(0.186)	-0.042(0.145)	7.7e-01
Breed:QH	-0.932(0.162)	-0.726(0.126)	**8.2e-09**
Breed:HR	-0.032(0.183)	-0.025(0.143)	8.6e-01
Breed:LR	-0.749(0.247)	-0.583(0.192)	**2.4e-03**
Group:obese$^-$laminitis$^-$	Reference	Reference	Reference
Group:obese$^-$laminitis$^+$	-0.075(0.102)	-0.058(0.079)	4.7e-01
Group:obese$^+$laminitis$^-$	0.561(0.094)	0.437(0.073)	**2.6e-09**
Group:obese$^+$laminitis$^+$	0.497(0.127)	0.387(0.099)	**9.4e-05**
+1 SD g Mcal/kg bwt per day	0.098(0.297)	0.076(0.232)	7.4e-01
+1 SD g CP/kg bwt per day	-0.083(0.133)	-0.065(0.103)	5.3e-01
+1 SD g NDF/kg bwt per day	-0.103(0.169)	-0.080(0.132)	5.4e-01
+1 SD g Starch/kg bwt per day	-0.077(0.075)	-0.060(0.059)	3.1e-01
+1 SD g WSC/kg bwt per day	-0.043(0.094)	-0.034(0.073)	6.5e-01
+1 SD hours grazing per day	0.007(0.058)	0.006(0.045)	9.0e-01
+1 SD hours exercise per day	-0.027(0.041)	-0.021(0.032)	5.1e-01
+1 SD hours stalled per day	0.024(0.041)	0.018(0.032)	5.7e-01
Month:DEC	Reference	Reference	Reference
Month:JAN	0.042(0.390)	0.033(0.303)	9.1e-01
Month:FEB	0.211(0.259)	0.164(0.202)	4.2e-01
Month:MAR	-0.138(0.238)	-0.108(0.185)	5.6e-01
Month:APR	-0.112(0.254)	-0.087(0.198)	6.6e-01
Month:MAY	0.409(0.244)	0.318(0.190)	9.4e-02
Month:JUN	-1.107(0.369)	-0.862(0.288)	**2.7e-03**
Month:JUL	-0.438(0.294)	-0.341(0.229)	1.4e-01
Month:AUG	-0.642(0.231)	-0.500(0.180)	**5.5e-03**
Month:SEP	-0.175(0.247)	-0.136(0.192)	4.8e-01
Month:OCT	0.505(0.246)	0.393(0.192)	**4.0e-02**
Month:NOV	0.144(0.322)	0.112(0.251)	6.5e-01
+1 SD latitude	0.016(0.060)	0.012(0.047)	7.9e-01
Oral thyroxine (no)	Reference	Reference	Reference
Oral thyroxine (yes)	0.129(0.189)	0.101(0.147)	4.9e-01
Researcher collected sample	Reference	Reference	Reference
Owner submitted sample	0.127(0.146)	0.099(0.114)	3.8e-01

Metabolic traits were mean centered and variance scaled to 1 standard deviation. Scaled fixed effect estimates represent the number of standard deviations change in the metabolic trait relative to the reference group.

Table A.15: Fixed effect estimates for fasting $\sqrt{}$adiponectin (ng/ml) from the full multivariate, multilevel model.

Parameter	scaled estimate(SE)	unscaled estimate(SE)	p-value
Intercept	0.469(0.168)	72.661(3.973)	
+1 SD age(yrs)	-0.035(0.038)	-0.820(0.891)	3.6e-01
Sex:gelding	Reference	Reference	Reference
Sex:mare	-0.055(0.082)	-1.298(1.943)	5.0e-01
Sex:stallion	0.135(0.181)	3.205(4.301)	4.6e-01
Breed:Morgan	Reference	Reference	Reference
Breed:Arab	-0.318(0.159)	-7.534(3.778)	**4.6e-02**
Breed:Pony	-0.043(0.135)	-1.029(3.189)	7.5e-01
Breed:TW	-0.275(0.179)	-6.526(4.232)	1.2e-01
Breed:QH	0.246(0.159)	5.832(3.765)	1.2e-01
Breed:HR	-0.077(0.178)	-1.819(4.227)	6.7e-01
Breed:LR	-0.106(0.249)	-2.512(5.903)	6.7e-01
Group:obese$^-$laminitis$^-$	Reference	Reference	Reference
Group:obese$^-$laminitis$^+$	-0.621(0.108)	-14.717(2.560)	**9.0e-09**
Group:obese$^+$laminitis$^-$	-0.138(0.099)	-3.277(2.356)	1.6e-01
Group:obese$^+$laminitis$^+$	-0.657(0.131)	-15.582(3.111)	**5.5e-07**
+1 SD g Mcal/kg bwt per day	-0.010(0.273)	-0.235(6.460)	9.7e-01
+1 SD g CP/kg bwt per day	0.036(0.118)	0.847(2.790)	7.6e-01
+1 SD g NDF/kg bwt per day	0.018(0.158)	0.422(3.738)	9.1e-01
+1 SD g Starch/kg bwt per day	0.131(0.071)	3.097(1.680)	6.5e-02
+1 SD g WSC/kg bwt per day	0.034(0.086)	0.815(2.029)	6.9e-01
+1 SD hours grazing per day	-0.070(0.056)	-1.670(1.324)	2.1e-01
+1 SD hours exercise per day	-0.074(0.041)	-1.749(0.982)	7.5e-02
+1 SD hours stalled per day	0.016(0.040)	0.369(0.960)	7.0e-01
Month:DEC	Reference	Reference	Reference
Month:JAN	-0.269(0.366)	-6.374(8.668)	4.6e-01
Month:FEB	-0.241(0.224)	-5.703(5.299)	2.8e-01
Month:MAR	0.295(0.197)	6.993(4.663)	1.3e-01
Month:APR	0.024(0.208)	0.563(4.930)	9.1e-01
Month:MAY	-0.396(0.208)	-9.376(4.921)	5.7e-02
Month:JUN	0.213(0.341)	5.061(8.080)	5.3e-01
Month:JUL	0.144(0.253)	3.410(6.000)	5.7e-01
Month:AUG	-0.196(0.196)	-4.657(4.653)	3.2e-01
Month:SEP	0.180(0.212)	4.265(5.020)	4.0e-01
Month:OCT	0.034(0.217)	0.817(5.152)	8.7e-01
Month:NOV	0.208(0.289)	4.920(6.841)	4.7e-01
+1 SD latitude	0.017(0.055)	0.413(1.306)	7.5e-01
Oral thyroxine (no)	Reference	Reference	Reference
Oral thyroxine (yes)	-0.243(0.191)	-5.770(4.537)	2.0e-01
Researcher collected sample	Reference	Reference	Reference
Owner submitted sample	-0.399(0.130)	-9.468(3.077)	**2.1e-03**

Metabolic traits were mean centered and variance scaled to 1 standard deviation. Scaled fixed effect estimates represent the number of standard deviations change in the metabolic trait relative to the reference group.

Appendix B

Supplemental Material: Chapter 3 Characterization of the equine response to an oral sugar challenge

Table B.1: *GCG* and *DPP4* PCR primers

Primer	Sequence (5' to 3')
DPP4 cDNA 1F	TGCAGACAGTGTGGAAGGTG
DPP4 cDNA 1R	CAAAGGCAGGAGCTGTGAAT
DPP4 cDNA 2F	GATCACGTGGAATGGGAAAG
DPP4 cDNA 2R	CAGCAATGAAGCAGGCTACA
DPP4 cDNA 3F	CGTGACATGGGTAACCACAG
DPP4 cDNA 3R	GGTATCAGATGATCTTGCCTCC
DPP4 cDNA 4F	TGCACAGCAGCAGGAATGAT
DPP4 cDNA 4R	AAGTGGCGTGTTCAAGTGTG
DPP4 cDNA 5F	CCTTGCAAGCACAGAAAACA
DPP4 cDNA 5R	CCACATGAGCCACTTCATACA
DPP4 exon 11F	AATCTTGAGCCATGGTGT
DPP4 exon 11R	CAGAGTGAATGGCTGGGAAT
GCG promoter/exon 1F	GAAGCCAGGCTGAAAAAGG
GCG promoter/exon 1R	GGTTTGGGGCACAGTTAAGT
GCG exon 2F	AAACATGGTTGCCTGTCTCC
GCG exon 2R	GCACAATGGTTAGAATACAGCC
GCG exon 3F	CCTTCCAGCTTCATGTGGTC
GCG exon 3R	GCTCATCAAATAGTGCCCAT
GCG exon 4F	GAATCCTGAACCCAAAGCTG
GCG exon 4R	GAAATAAGTCCCATTGGTGGG
GCG exon 5F	CGTTTTCTCAATGACCCTCAC
GCG exon 5R	GTTCCCAATGTGAGGGAGAG
GCG exon 6F	TTGACAAGTCCACAAGACTAAGG
GCG exon 6R	GATATGGAAGGACTGGCAGC
GUE F	GACACTCAAACGAAGCCAGG
GUE R	ACTGGCAGGTCACAAAGCTC

Figure B.1: Proglucagon upstream enhancer, promoter region, and exon 1 cross species nucleotide alignment

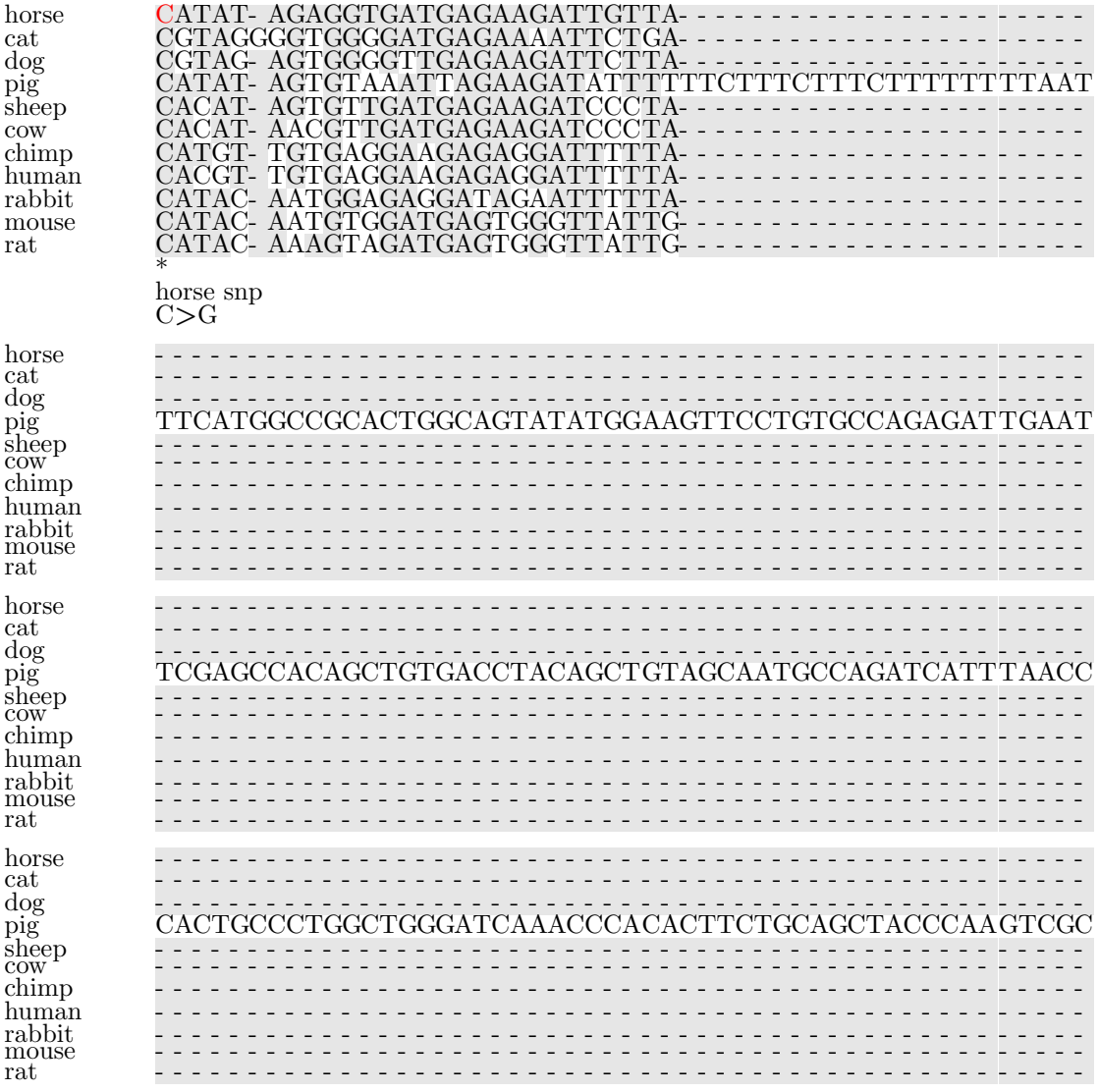

horse CATAT- AGAGGTGATGAGAAGATTGTTA- - - -

Figure B.1 *Continued from previous page*

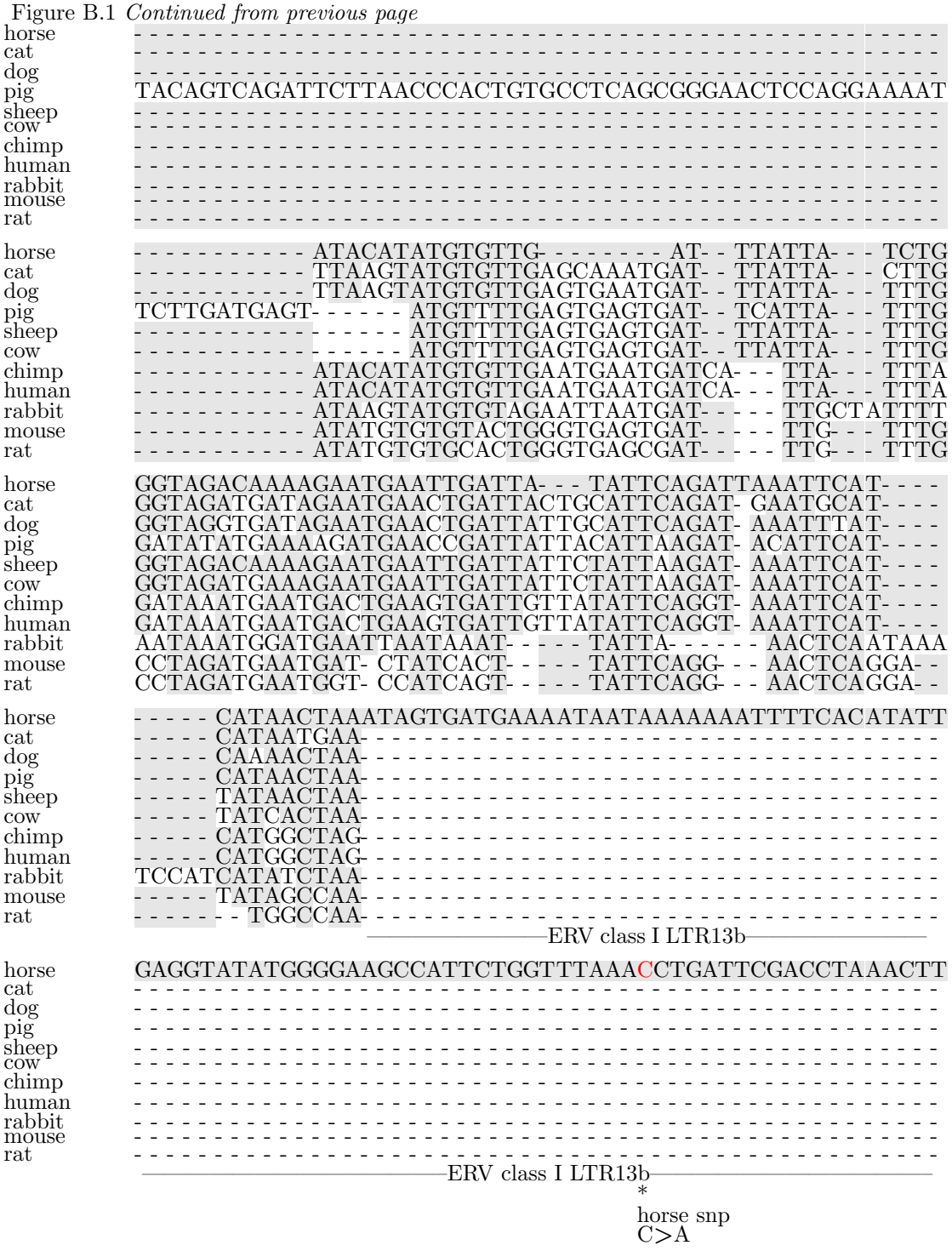


Figure B.1 *Continued on next page*

Figure B.1 *Continued from previous page*

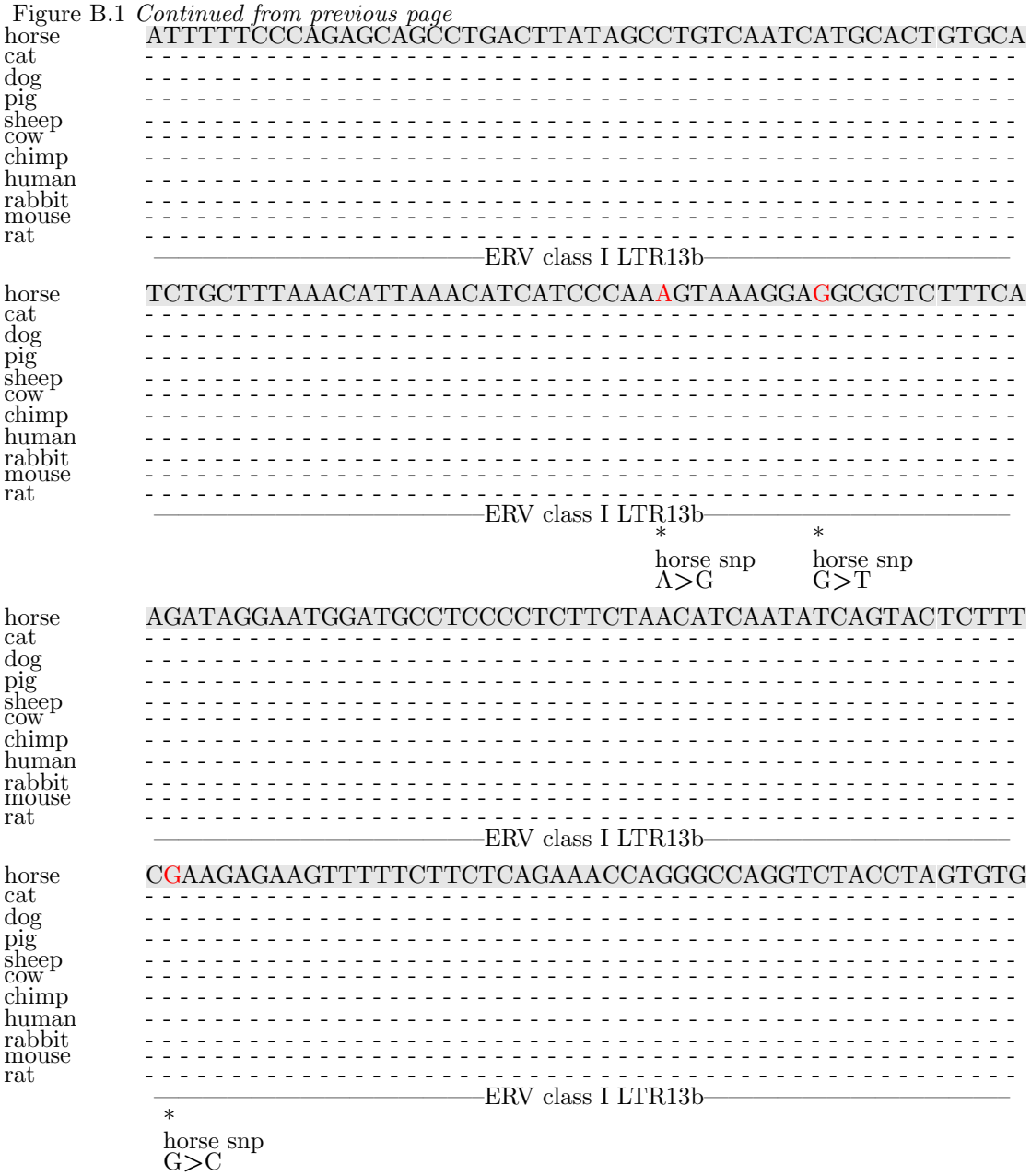


Figure B.1 *Continued on next page*

Figure B.1 *Continued from previous page*

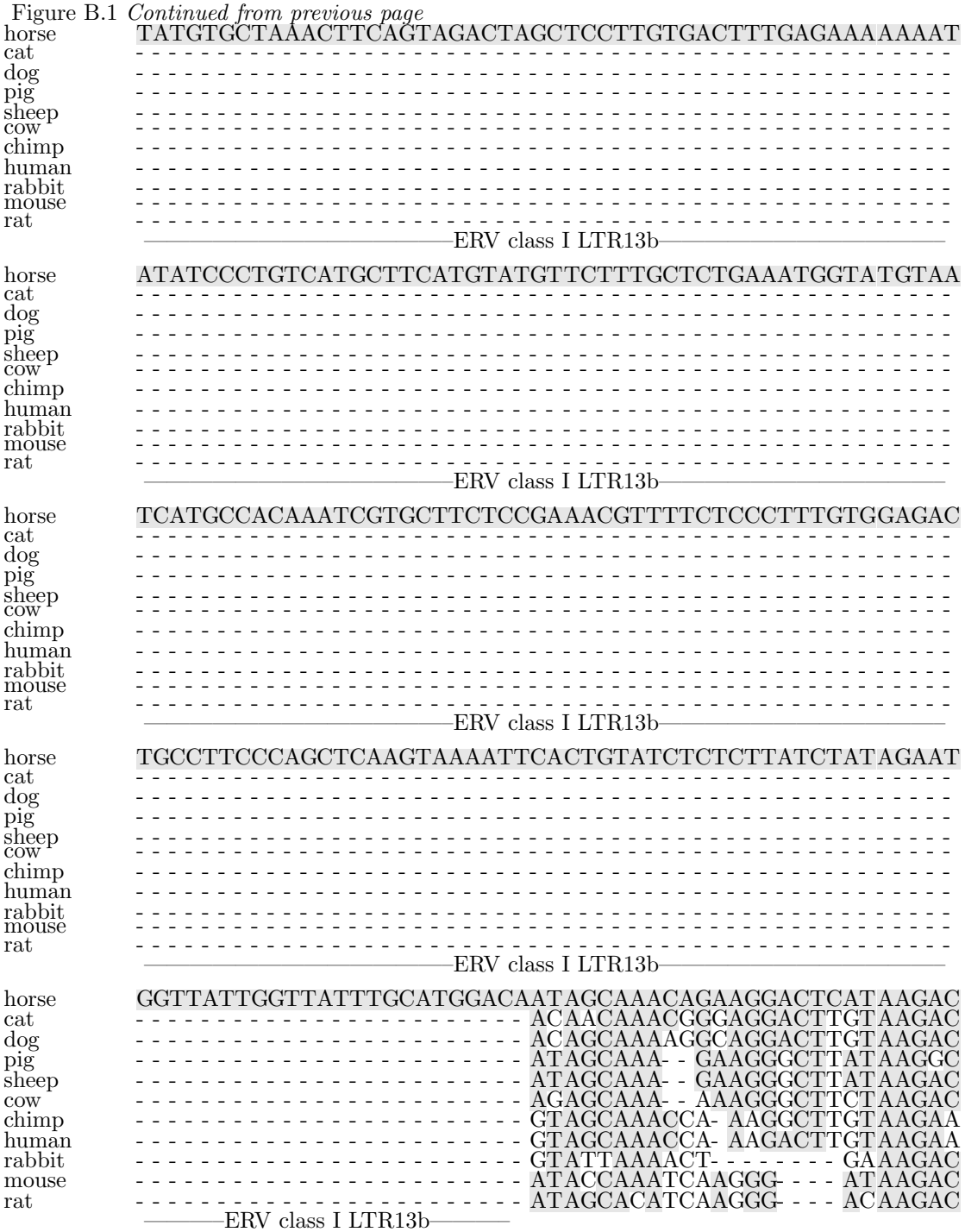


Figure B.1 *Continued on next page*

Figure B.1 *Continued from previous page*

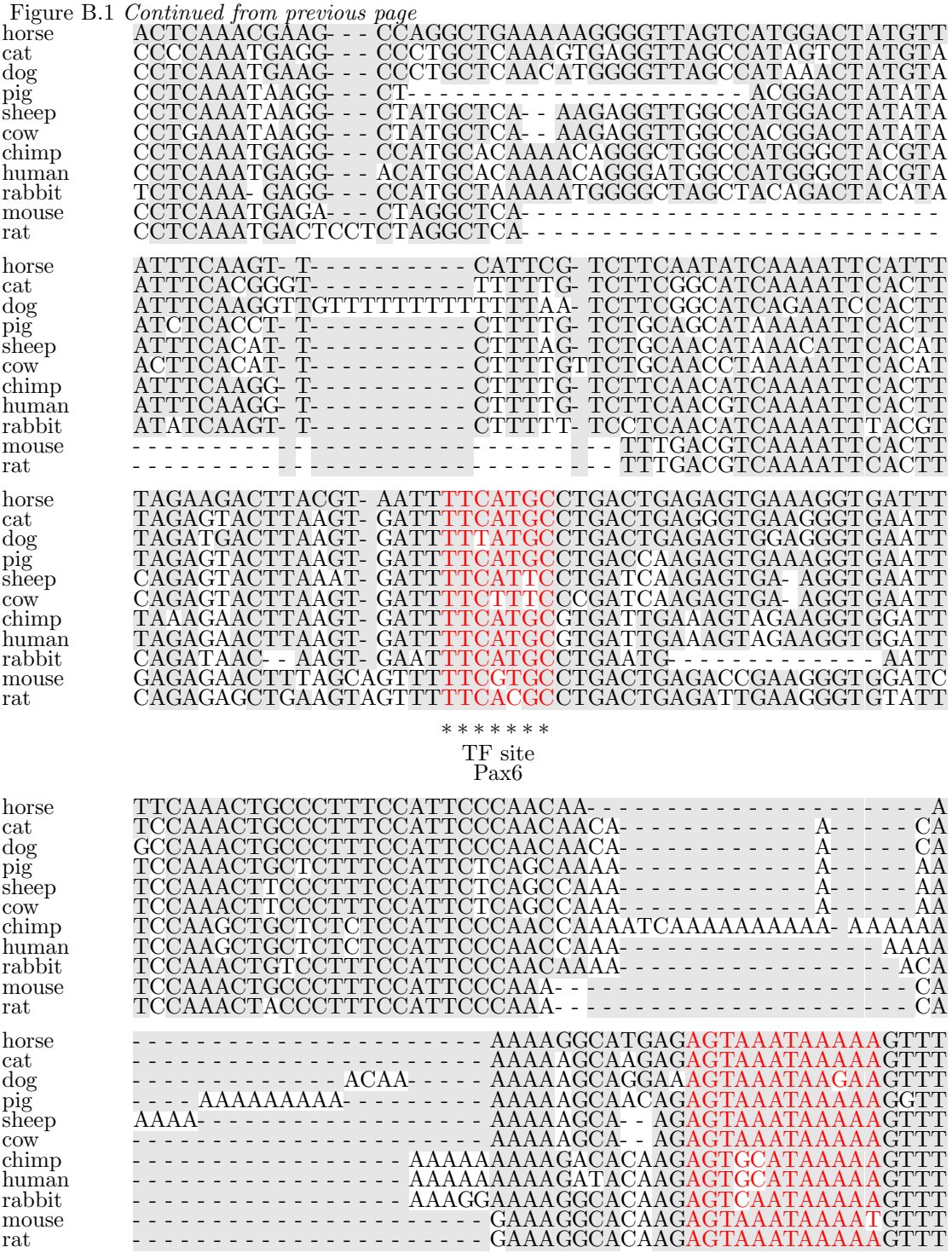


Figure B.1 *Continued on next page*

Figure B.1 *Continued from previous page*

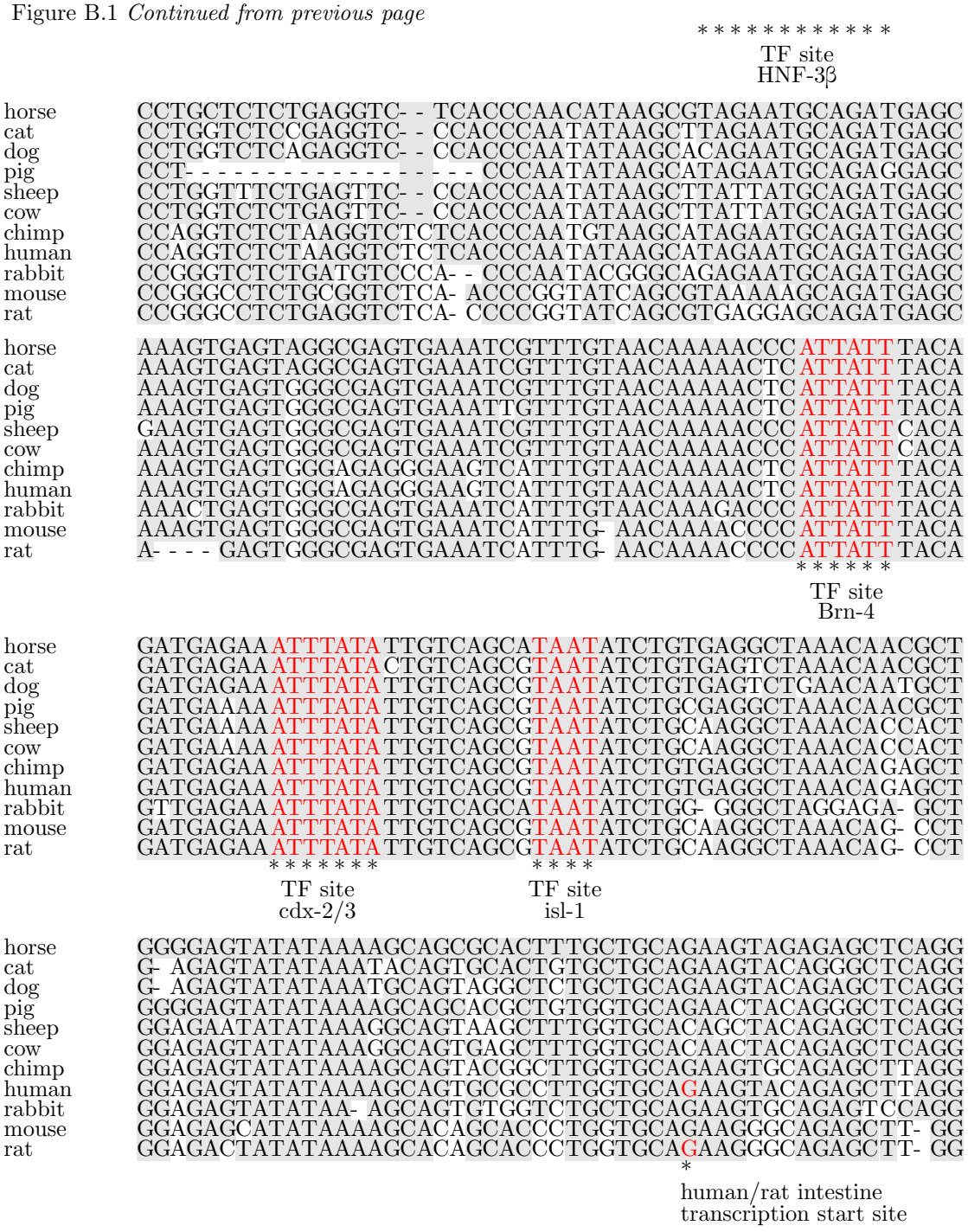


Figure B.1 *Continued on next page*

Figure B.1 *Continued from previous page*

```
horse   AGACA- - GCACATCCAAAGTTCTGAAAGGAGGGCGCCTCTCTCTCCACC
cat     ACACA- - - CACATCCAAAGTTCCGAAAAGAGGGCTCGCTCTCTCTCCACC
dog     ACACA- - GCACAGCCAAAGTTCCGAAAAGAGGGCTCGCTCTCTCCCCACC
pig     ACACT- - GCACACACGAAGTTCTGA- AAAGAGTCTCACTCTCTTTCCACC
sheep   ACACT- - GCACACCCAA- - - - - - - - - ACGAGGGCTCACTCTCTCTCCACC
cow     ACACT- - GCACACCCAA- - - - - - - - - ACGAGGGCTCACTCTCTCTTCACC
chimp   ACACAGAGCACATCAAAAGTTCCCA- AAGAGGGCTTGCTCTCTCTTCACC
human   ACACAGAGCACATCAAAAGTTCCCA- AAGAGGGCTTGCTCTCTCTTCACC
rabbit  GCACAGAGCACATCCAGTGTCCACA- - - GAGGGCTTGCTCTCTCTCCACC
mouse   GCCCAGGACACACTCAAAGTTCCCA- A- GGGGACTCCCTCTGTCTACACC
rat     GCGCAGAACACACTCAAAGTTCCCA- A- AGGAGCTCCACCTGTCTACACC
                                                  *
                                                  horse snp
                                                  C>G

horse   TGCTCTGCTCCACATCACAGT- GTCAGAAG
cat     TGCTCTGTTCCA- GTC- AAGT- GTCAGAAG
dog     TGCTCTGTTCCA- ATCGCGGT- GTCAGAAG
pig     TGCCCTGTTCCACCTTCTGGT- GTCAAAAG
sheep   TGCTCTGTTCCACCTCCTGGT- GTCAGAAG
cow     TGCTCTGTTCCACCTCCTGGT- GTCAGAAG
chimp   TGCTCTGTTCTACAGCACACT- ACCAGAAG
human   TGCTCTGTTCTACAGCACACT- ACCAGAAG
rabbit  TGCTCTGTTCCACACGCCAGTACTCAGAAG
mouse   TGTTCG- - - - - - CAGCTCAGG- CTCACAAG
rat     TCCTCT- - - - - - CAGCTCAGT- CCCACAAG
                                    *
                                    human/rat
                                    final nucleotide non-
                                    coding exon 1
```

Horse nucleotide sequence=reverse complement of Ensembl generated EquCab2.0 assembly chr18:42522863-42523969. Cross species alignments generated with Ensembl.[255] Putative transcription factor (TF) binding sites obtained from Nian et al 1999[171] and transcription start site obtained from Novak et al 1987.[256] Identification of horse endogenous retrovirus class I long terminal repeat 13b (ERV class I LTR13b) obtained using RepeatMasker.[172]

Table B.2: Glucose mg/dl area under the curve and trajectory fixed effect estimate(se)

Univariate predictor	AUC	Trajectory		
		intercept	linear slope	quadratic slope
Ref group (mean age)	11417.12(252.38)	74.10(0.89)	10.67(0.56)	-0.98(0.07)
1 SD increase in age	-349.42(134.50)**	-1.25(0.72)	-1.27(0.56)*	0.14(0.07)*
Ref group (gelding)	11313.48(413.53)	74.43(1.83)	9.75(1.30)	-0.89(0.15)
mare	97.47(355.15)	-0.68(1.87)	1.17(1.46)	-0.10(0.17)
stallion	289.10(707.92)	2.24(3.81)	0.51(3.03)	-0.19(0.36)
Ref group (Morgan breed)	11035.56(371.09)	74.35(1.48)	8.74(0.77)	-0.77(0.09)
Welsh Pony	824.63(555.46)	-0.39(2.18)	3.86(1.09)***	-0.41(0.13)**
Ref group (non-obese, no lam)	11284.10(310.79)	73.32(1.12)	9.41(0.80)	-0.84(0.09)
non-obese, prior laminitis	-720.99(384.73)	-2.44(1.92)	0.04(1.48)	0.00(0.17)
obese, no prior laminitis	988.58(381.17)**	5.01(2.03)*	2.65(1.60)	-0.26(0.19)
obese, prior laminitis	508.64(449.68)	2.80(2.07)	4.87(1.57)**	-0.56(0.18)**
Ref group(mean Mcal)	11440.93(253.30)	74.12(1.02)	10.67(0.56)	-0.98(0.07)
1 SD increase in Mcal	-380.35(187.59)*	0.10(0.87)	-1.44(0.56)*	0.12(0.07)
Ref group (mean nutrients)	11462.49(256.36)	74.35(1.17)	10.67(0.53)	-0.98(0.06)
1 SD increase in Starch	-228.87(258.30)	0.45(1.27)	-1.67(0.76)*	0.20(0.09)*
1 SD increase in CP	0.88(253.25)	-0.13(1.25)	0.23(0.76)	-0.04(0.09)
1 SD increase in NDF	-637.61(341.97)	-1.76(1.63)	-3.21(0.92)***	0.36(0.11)***
1 SD increase in WSC	374.37(348.35)	2.30(1.66)	2.32(0.93)*	-0.32(0.11)**
Ref group (mean triglycerides)	11393.58(314.44)	74.03(1.15)	10.67(0.57)	-0.98(0.07)
1 SD increase in triglycerides	-100.65(163.38)	-1.53(0.81)	1.00(0.57)	-0.08(0.07)
Ref group (mean NEFA)	11428.23(200.16)	74.16(1.00)	10.67(0.55)	-0.98(0.07)
1 SD increase in NEFA	386.43(150.73)*	-0.27(0.81)	1.89(0.55)***	-0.16(0.07)*
Ref group (mean leptin)	11385.18(323.23)	74.08(1.08)	10.67(0.57)	-0.98(0.07)
1 SD increase in leptin	246.06(135.33)	0.19(0.74)	0.64(0.57)	-0.07(0.07)
Ref group (mean adiponectin)	11400.07(294.23)	74.09(0.99)	10.67(0.57)	-0.98(0.07)
1 SD increase in adiponectin	29.49(144.38)	-0.14(0.75)	-0.49(0.58)	0.08(0.07)
Ref group (mean DPP4 activity)	11397.33(276.57)	74.08(1.01)	10.67(0.56)	-0.98(0.07)
1 SD increase in DPP4 activity	197.73(164.01)	0.16(0.80)	1.52(0.56)**	-0.17(0.07)*
Ref group (*DPP4* 0 copies)	11371.27(337.26)	74.92(1.33)	10.19(0.90)	-0.95(0.11)
DPP4 1 copy minor allele	32.27(189.49)	-0.93(1.01)	0.54(0.80)	-0.03(0.09)

Table B.2 *Continued on next page*

Table B.2 *Continued from previous page*

Univariate predictor	AUC	Trajectory		
		intercept	linear slope	quadratic slope
Ref group (*GCG* haplotype A)	11513.16(435.94)	72.75(1.89)	11.09(1.26)	-0.95(0.15)
GCG haplotype B	99.05(267.83)	2.24(1.34)	-1.19(1.00)	0.11(0.12)
GCG haplotype C	-185.83(252.70)	-0.73(1.28)	0.70(0.96)	-0.13(0.11)
GCG haplotype D	-645.89(476.14)	-0.55(2.47)	0.01(1.90)	-0.21(0.22)
GCG haplotype E	-20.67(362.77)	2.55(1.90)	-0.89(1.47)	0.05(0.17)

Asterisks indicate significance of beta estimates (*<0.05, **<0.01, ***<0.001). AUC estimates were calculated in minutes. Time was coded in 15 minute intervals and centered at zero minutes for the trajectory estimates. Results obtained from a multivariate response model (responses included glucose, insulin, active GLP-1 and total GLP-1 measured at 7 time points over a 2 hour period). *GCG* haplotype A=highest frequency haplotype, *GCG* haplotype E=pooled rare haplotypes with frequency $<$ **0.05**

Table B.3: Insulin μIU/ml area under the curve and trajectory fixed effect estimate(se)

Univariate predictor	AUC	Trajectory		
		intercept	linear slope	quadratic slope
Ref group (mean age)	3526.22(537.69)	6.16(1.07)	11.74(1.12)	-1.06(0.12)
1 SD increase in age	-8.45(247.87)	1.24(0.62)*	-0.07(1.12)	-0.01(0.12)
Ref group (gelding)	2902.20(729.54)	4.61(1.69)	7.43(2.49)	-0.65(0.27)
mare	757.26(629.65)	1.79(1.60)	5.35(2.79)	-0.50(0.31)
stallion	1104.12(1267.99)	2.94(3.25)	5.40(5.79)	-0.63(0.64)
Ref group (Morgan breed)	2877.62(655.87)	4.98(1.36)	8.26(1.51)	-0.71(0.17)
Welsh Pony	1474.56(981.09)	2.66(2.02)	6.95(2.14)**	-0.70(0.24)**
Ref group (non-obese, no lam)	2991.86(536.77)	5.11(1.01)	9.15(1.57)	-0.79(0.17)
non-obese, prior laminitis	1100.66(706.11)	0.54(1.68)	5.05(2.90)	-0.45(0.32)
obese, no prior laminitis	390.10(716.33)	0.19(1.76)	0.51(3.11)	-0.05(0.34)
obese, prior laminitis	2124.56(804.10)**	6.91(1.84)***	8.89(3.05)**	-1.04(0.33)**
Ref group(mean Mcal)	3563.08(534.22)	6.07(1.16)	11.74(1.11)	-1.06(0.12)
1 SD increase in Mcal	-313.00(355.82)	0.24(0.82)	-1.78(1.11)	0.16(0.12)

Table B.3 *Continued on next page*

Table B.3 *Continued from previous page*

Univariate predictor	AUC	Trajectory		
		intercept	linear slope	quadratic slope
Ref group (mean nutrients)	3640.25(243.34)	6.27(1.03)	11.74(1.03)	-1.06(0.11)
1 SD increase in Starch	-531.69(349.22)	-0.21(1.07)	-2.05(1.47)	0.18(0.16)
1 SD increase in CP	627.25(347.10)	0.63(1.06)	2.47(1.46)	-0.28(0.16)
1 SD increase in NDF	-730.15(421.01)	-1.18(1.41)	-2.70(1.78)	0.27(0.20)
1 SD increase in WSC	32.86(423.75)	1.02(1.42)	0.08(1.79)	-0.03(0.20)
Ref group (mean triglycerides)	3564.52(475.35)	6.23(0.96)	11.74(1.09)	-1.06(0.12)
1 SD increase in triglycerides	463.22(278.13)	-0.03(0.68)	2.47(1.09)*	-0.16(0.12)
Ref group (mean NEFA)	3521.08(534.73)	6.20(0.91)	11.74(1.11)	-1.06(0.12)
1 SD increase in NEFA	18.55(291.54)	1.26(0.68)	1.43(1.11)	-0.15(0.12)
Ref group (mean leptin)	3511.34(528.02)	6.16(1.09)	11.74(1.07)	-1.06(0.12)
1 SD increase in leptin	881.39(233.76)***	0.31(0.63)	3.20(1.08)**	-0.27(0.12)*
Ref group (mean adiponectin)	3508.46(476.93)	6.15(1.08)	11.74(1.06)	-1.06(0.12)
1 SD increase in adiponectin	-735.02(246.82)**	0.26(0.64)	-3.79(1.06)***	0.33(0.12)**
Ref group (mean DPP4 activity)	3512.55(517.39)	6.10(1.04)	11.74(1.09)	-1.06(0.12)
1 SD increase in DPP4 activity	307.07(289.48)	1.26(0.69)	2.68(1.09)*	-0.31(0.12)**
Ref group (*DPP4* 0 copies)	3377.04(608.34)	5.15(1.31)	9.56(1.74)	-0.76(0.19)
DPP4 1 copy minor allele	164.92(340.29)	1.14(0.86)	2.48(1.53)	-0.35(0.17)*
Ref group (*GCG* haplotype A)	3728.59(763.73)	5.28(1.67)	12.04(2.47)	-1.04(0.27)
GCG haplotype B	-53.04(473.38)	1.86(1.14)	-1.23(1.95)	0.08(0.21)
GCG haplotype C	-245.52(447.63)	0.47(1.09)	0.75(1.87)	-0.07(0.20)
GCG haplotype D	295.95(853.69)	-0.29(2.10)	3.02(3.70)	-0.50(0.40)
GCG haplotype E	-415.28(653.02)	-1.81(1.62)	-2.22(2.87)	0.20(0.31)

Asterisks indicate significance of beta estimates (*<0.05, **<0.01, ***<0.001). AUC estimates were calculated in minutes. Time was coded in 15 minute intervals and centered at zero minutes for the trajectory estimates. Results obtained from a multivariate response model (responses included glucose, insulin, active GLP-1 and total GLP-1 measured at 7 time points over a 2 hour period).

Table B.4: Active GLP-1 pM area under the curve and trajectory fixed effect estimate(se)

Univariate predictor	AUC	Trajectory		
		intercept	linear slope	quadratic slope
Ref group (mean age)	1377.16(221.86)	7.31(1.52)	2.32(0.28)	-0.26(0.03)
1 SD increase in age	41.04(132.80)	0.55(0.97)	-0.17(0.28)	0.02(0.03)
Ref group (gelding)	1333.04(351.12)	4.91(2.65)	2.70(0.61)	-0.24(0.06)
mare	67.20(340.28)	2.30(2.50)	-0.30(0.69)	-0.04(0.07)
stallion	-60.61(687.86)	9.17(5.03)	-3.27(1.42)*	0.23(0.15)
Ref group (Morgan breed)	1536.88(298.64)	8.80(2.22)	2.38(0.39)	-0.26(0.04)
Welsh Pony	-379.17(443.00)	-3.18(3.29)	-0.11(0.55)	0.00(0.06)
Ref group (non-obese, no lam)	1464.45(258.74)	9.36(1.97)	2.06(0.39)	-0.23(0.04)
non-obese, prior laminitis	-295.17(378.01)	-3.99(2.80)	-0.02(0.71)	-0.01(0.08)
obese, no prior laminitis	251.88(385.87)	-2.12(2.84)	2.18(0.77)**	-0.21(0.08)*
obese, prior laminitis	-634.42(427.78)	-6.38(3.19)*	-0.52(0.75)	0.04(0.08)
Ref group(mean Mcal)	1364.58(212.14)	7.25(1.59)	2.32(0.28)	-0.26(0.03)
1 SD increase in Mcal	131.16(165.17)	1.64(1.27)	0.01(0.28)	0.00(0.03)
Ref group (mean nutrients)	1364.91(127.51)	7.26(1.56)	2.32(0.28)	-0.26(0.03)
1 SD increase in Starch	29.26(182.99)	1.43(1.61)	-0.53(0.40)	0.03(0.04)
1 SD increase in CP	-294.08(181.87)	-2.25(1.59)	-0.43(0.39)	0.01(0.04)
1 SD increase in NDF	-43.42(220.61)	0.06(2.12)	-0.23(0.48)	0.02(0.05)
1 SD increase in WSC	447.46(222.04)*	3.74(2.13)	0.39(0.48)	-0.04(0.05)
Ref group (mean triglycerides)	1371.64(217.67)	7.28(1.53)	2.32(0.28)	-0.26(0.03)
1 SD increase in triglycerides	-55.89(148.80)	-1.56(1.09)	0.08(0.28)	-0.02(0.03)
Ref group (mean NEFA)	1358.47(209.18)	7.27(1.48)	2.32(0.28)	-0.26(0.03)
1 SD increase in NEFA	-234.00(146.38)	-1.93(1.09)	-0.21(0.28)	0.02(0.03)
Ref group (mean leptin)	1380.04(220.38)	7.32(1.51)	2.32(0.28)	-0.26(0.03)
1 SD increase in leptin	51.66(132.68)	-0.08(0.98)	0.26(0.28)	-0.02(0.03)
Ref group (mean adiponectin)	1383.98(208.24)	7.36(1.44)	2.32(0.28)	-0.26(0.03)
1 SD increase in adiponectin	196.36(136.04)	1.60(1.00)	-0.06(0.28)	0.02(0.03)
Ref group (mean DPP4 activity)	1378.22(198.80)	7.24(1.58)	2.32(0.28)	-0.26(0.03)
1 SD increase in DPP4 activity	-183.79(146.34)	-0.71(1.12)	-0.18(0.28)	0.01(0.03)

Table B.4 *Continued on next page*

Table B.4 *Continued from previous page*

Univariate predictor	AUC	Trajectory		
		intercept	linear slope	quadratic slope
Ref group (*DPP4* 0 copies)	1378.27(275.31)	7.60(1.93)	2.42(0.43)	-0.28(0.05)
DPP4 1 copy minor allele	-2.95(182.93)	-0.35(1.35)	-0.11(0.38)	0.02(0.04)
Ref group (*GCG* haplotype A)	1262.47(345.63)	4.85(2.43)	2.97(0.60)	-0.29(0.06)
GCG haplotype B	-79.35(243.08)	1.63(1.79)	-0.86(0.48)	0.06(0.05)
GCG haplotype C	141.75(232.49)	0.77(1.72)	-0.31(0.46)	0.01(0.05)
GCG haplotype D	352.84(448.35)	3.05(3.30)	0.57(0.90)	-0.10(0.10)
GCG haplotype E	280.91(345.26)	4.88(2.54)	-0.48(0.70)	0.02(0.07)

Asterisks indicate significance of beta estimates (*<0.05, **<0.01, ***<0.001). AUC estimates were calculated in minutes. Time was coded in 15 minute intervals and centered at zero minutes for the trajectory estimates. Results obtained from a multivariate response model (responses included glucose, insulin, active GLP-1 and total GLP-1 measured at 7 time points over a 2 hour period). *GCG* haplotype A=highest frequency haplotype, *GCG* haplotype E=pooled rare haplotypes with frequency $<$ **0.05**

Table B.5: Total GLP-1 pM area under the curve and trajectory fixed effect estimate(se)

Univariate predictor	AUC	Trajectory		
		intercept	linear slope	quadratic slope
Ref group (mean age)	2430.76(421.09)	16.94(3.12)	2.04(0.34)	-0.25(0.03)
1 SD increase in age	89.98(230.12)	0.79(1.85)	-0.24(0.34)	0.04(0.03)
Ref group (gelding)	2066.28(634.84)	9.89(4.91)	3.39(0.74)	-0.33(0.07)
mare	443.67(587.35)	7.51(4.64)	-1.42(0.83)	0.08(0.08)
stallion	788.41(1191.32)	22.30(9.34)*	-5.84(1.73)***	0.44(0.17)**
Ref group (Morgan breed)	2797.17(539.02)	19.22(4.26)	2.53(0.48)	-0.29(0.05)
Welsh Pony	-864.40(801.39)	-5.40(6.33)	-0.98(0.68)	0.07(0.07)
Ref group (non-obese, no lam)	2819.31(451.16)	21.21(3.72)	2.02(0.47)	-0.24(0.05)
non-obese, prior laminitis	-1086.52(644.91)	-10.48(5.27)*	-0.34(0.86)	0.01(0.08)
obese, no prior laminitis	-146.40(664.49)	-7.08(5.36)	2.75(0.93)**	-0.24(0.09)**
obese, prior laminitis	-1339.43(721.11)	-10.20(5.99)	-2.03(0.91)*	0.17(0.09)

Table B.5 *Continued on next page*

Table B.5 *Continued from previous page*

Univariate predictor	AUC	Trajectory		
		intercept	linear slope	quadratic slope
Ref group(mean Mcal)	2402.46(415.35)	16.97(3.26)	2.04(0.34)	-0.25(0.03)
1 SD increase in Mcal	86.41(298.70)	-0.83(2.47)	0.82(0.34)*	-0.08(0.03)**
Ref group (mean nutrients)	2363.22(222.69)	16.56(2.23)	2.04(0.33)	-0.25(0.03)
1 SD increase in Starch	254.28(319.59)	5.12(2.75)	-1.05(0.47)*	0.06(0.05)
1 SD increase in CP	-729.35(317.64)*	-3.72(2.73)	-1.02(0.46)*	0.03(0.05)
1 SD increase in NDF	352.12(385.28)	2.56(3.43)	0.23(0.56)	-0.02(0.06)
1 SD increase in WSC	161.31(387.79)	0.04(3.46)	0.92(0.57)	-0.08(0.06)
Ref group (mean triglycerides)	2413.86(391.24)	16.72(2.92)	2.04(0.34)	-0.25(0.03)
1 SD increase in triglycerides	-223.01(252.03)	-1.46(2.08)	-0.71(0.34)*	0.05(0.03)
Ref group (mean NEFA)	2405.93(370.96)	16.80(2.68)	2.04(0.34)	-0.25(0.03)
1 SD increase in NEFA	-436.79(248.57)	-2.98(2.06)	-0.63(0.34)	0.04(0.03)
Ref group (mean leptin)	2440.56(420.92)	16.96(3.12)	2.04(0.34)	-0.25(0.03)
1 SD increase in leptin	34.48(230.06)	-0.98(1.85)	0.66(0.34)	-0.06(0.03)
Ref group (mean adiponectin)	2447.49(385.32)	17.04(2.92)	2.04(0.34)	-0.25(0.03)
1 SD increase in adiponectin	320.24(235.27)	1.41(1.91)	0.38(0.34)	-0.02(0.03)
Ref group (mean DPP4 activity)	2441.65(385.32)	16.89(3.09)	2.04(0.34)	-0.25(0.03)
1 SD increase in DPP4 activity	-216.44(254.85)	-0.12(2.14)	-0.47(0.34)	0.03(0.03)
Ref group (*DPP4* 0 copies)	2611.37(498.50)	18.88(3.79)	1.98(0.54)	-0.25(0.05)
DPP4 1 copy minor allele	-202.48(316.61)	-2.23(2.55)	0.08(0.48)	0.00(0.05)
Ref group (*GCG* haplotype A)	2055.48(635.51)	10.29(4.80)	3.33(0.74)	-0.32(0.07)
GCG haplotype B	194.23(424.51)	5.37(3.40)	-1.40(0.58)*	0.10(0.06)
GCG haplotype C	258.81(404.03)	3.17(3.24)	-0.91(0.56)	0.05(0.05)
GCG haplotype D	171.80(780.58)	0.93(6.21)	0.96(1.11)	-0.12(0.11)
GCG haplotype E	679.67(600.37)	10.68(4.77)*	-0.82(0.86)	0.00(0.08)

Asterisks indicate significance of beta estimates (*<0.05, **<0.01, ***<0.001). AUC estimates were calculated in minutes. Time was coded in 15 minute intervals and centered at zero minutes for the trajectory estimates. Results obtained from a multivariate response model (responses included glucose, insulin, active GLP-1 and total GLP-1 measured at 7 time points over a 2 hour period). *GCG* haplotype A=highest frequency haplotype, *GCG* haplotype E=pooled rare haplotypes with frequency $< \mathbf{0.05}$

Factors with a p-value significance level $< \mathbf{0.2}$ in the univariate predictor analysis were included in a multiple predictor regression analysis of glucose, insulin, and GLP-1 trajectories. In the multiple regression analysis glucose trajectories were significantly associated with clinical group, caloric intake, diet nutrient composition, and NEFA blood concentrations (Table B.6). The different clinical group glucose trajectories differed at the intercept (measurement at time zero) with obese, no prior laminitis horses having a higher baseline glucose measurement. Insulin trajectories were associated with leptin, triglycerides, and *DPP4* genotype (Table B.7). Horses with higher leptin levels also exhibited a higher initial rate of change in insulin and subsequently a faster deceleration in the initial rate of change in attempt to prevent a delayed return to baseline insulin levels. DPP4 genotype was also positively associated with insulin initial rate of change and deceleration rate. GLP-1 active and total responses were less dynamic in stallions (Tables B.8,B.9). Total GLP-1 trajectories were also associated with clinical group, dietary nutrient composition, and haplotype at the *GCG* upstream enhancer region (Table B.9). Obese horses with no prior laminitis exhibited a higher initial rate of change in total GLP-1 and subsequently a faster deceleration in the initial rate of change in attempt to prevent a delayed return to baseline levels.

Table B.6: Glucose mg/dl trajectory multiple regression association tests

parameter	intercept estimate(se)	time estimate(se)	time^2 estimate(se)	overall signficance (wald χ^2, df) pval
mean trajectory	72.88(2.50)	9.38(1.22)	-0.78(0.14)	
age	-1.06(0.74)	-0.51(0.51)	0.05(0.06)	(3.7, 3df)=3.0E-01
breed				
Morgan	Reference	Reference	Reference	
Welsh Pony	0.74(3.83)	2.82(1.80)	-0.34(0.21)	(2.8, 3df)=4.3E-01
clinical group				
obese$^-$lam$^-$	Reference	Reference	Reference	
obese$^-$lam$^+$	-1.92(2.31)	-2.04(1.65)	0.15(0.19)	
obese$^+$lam$^-$	4.91(2.22)*	1.46(1.61)	-0.16(0.19)	
obese$^+$lam$^+$	3.32(2.70)	0.24(1.91)	-0.16(0.23)	**(20.0, 9df)=1.8E-02**
Mcal	4.83(5.19)	-8.13(2.84)**	0.85(0.32)**	**(8.6, 3df)=3.6E-02**
nutrients				
CP	-0.99(2.45)	2.80(1.31)*	-0.34(0.15)*	
NDF	-5.59(5.48)	6.54(3.09)*	-0.64(0.35)	
WSC	0.93(1.95)	0.79(0.99)	-0.19(0.12)	
Starch	0.33(1.88)	0.41(1.04)	-0.02(0.12)	**(21.4, 12df)=4.5E-02**
DPP4 activity	-0.10(0.96)	0.64(0.65)	-0.08(0.08)	(1.1, 3df)=7.8E-01
NEFA	-0.62(0.97)	0.76(0.66)	0.00(0.08)	**(10.1, 3df)=1.7E-02**
triglycerides	-0.84(0.90)	-0.77(0.65)	0.13(0.08)	(5.2, 3df)=1.6E-01

Asterisks indicate significance of beta estimates (*<0.05, **<0.01, ***<0.001). Results obtained from a multivariate response model (responses included glucose, insulin, active GLP-1 and total GLP-1 measured at 7 time points over a 2 hour period). Time was coded in 15 minute intervals and centered at zero minutes for the trajectory estimates. Continuous predictors are scaled to zero mean and 1 standard deviation. obese$^-$lam$^-$=non-obese/no prior laminitis, obese$^-$lam$^+$=non-obese/prior laminitis, obese$^+$lam$^-$=obese/no prior laminitis, obese$^+$lam$^+$=obese/prior laminitis

Table B.7: Insulin μIU/ml trajectory multiple regression association tests

parameter	intercept estimate(se)	time estimate(se)	time^2 estimate(se)	overall signficance (wald χ^2, df) pval
mean trajectory	3.59(3.43)	6.96(2.67)	-0.43(0.29)	
breed				
Morgan	Reference	Reference	Reference	
Welsh Pony	-0.10(5.18)	0.92(3.51)	-0.09(0.38)	(0.1, 3df)=1.0E+00
clinical group				
obese$^-$lam$^-$	Reference	Reference	Reference	
obese$^-$lam$^+$	1.28(2.00)	3.72(3.01)	-0.40(0.34)	
obese$^+$lam$^-$	1.11(1.93)	1.11(3.08)	-0.14(0.34)	
obese$^+$lam$^+$	7.10(2.37)**	5.88(3.35)	-0.86(0.37)*	(16.6, 9df)=5.5E-02
nutrients				
CP	-0.91(1.69)	1.81(1.66)	-0.24(0.19)	
NDF	-0.68(2.60)	-3.16(2.15)	0.30(0.24)	
WSC	1.92(2.27)	1.33(1.82)	-0.15(0.20)	
Starch	0.30(1.62)	-0.91(1.66)	0.06(0.18)	(6.8, 12df)=8.7E-01
DPP4 activity	1.21(0.88)	0.26(1.30)	-0.06(0.14)	(2.1, 3df)=5.5E-01
adiponectin	-0.30(0.76)	-2.31(1.14)*	0.20(0.12)	(4.9, 3df)=1.8E-01
leptin	0.30(0.64)	2.88(0.94)**	-0.25(0.10)*	**(10.8, 3df)=1.3E-02**
NEFA	1.36(0.88)	-0.97(1.32)	0.08(0.14)	(2.7, 3df)=4.4E-01
triglycerides	-0.80(0.82)	-0.78(1.27)	0.22(0.14)	**(8.6, 3df)=3.5E-02**
DPP4 genotype	1.31(0.86)	2.76(1.29)*	-0.39(0.14)**	**(11.8, 3df)=8.1E-03**

Asterisks indicate significance of beta estimates (*<0.05, **<0.01, ***<0.001). Results obtained from a multivariate response model (responses included glucose, insulin, active GLP-1 and total GLP-1 measured at 7 time points over a 2 hour period). Time was coded in 15 minute intervals and centered at zero minutes for the trajectory estimates. Continuous predictors are scaled to zero mean and 1 standard deviation. obese$^-$lam$^-$=non-obese/no prior laminitis, obese$^-$lam$^+$=non-obese/prior laminitis, obese$^+$lam$^-$=obese/no prior laminitis, obese$^+$lam$^+$=obese/prior laminitis

Table B.8: Active GLP-1 pM trajectory multiple regression association tests

parameter	intercept estimate(se)	time estimate(se)	time^2 estimate(se)	overall signficance (wald χ^2, df) pval
mean trajectory	4.58(3.08)	2.73(0.77)	-0.25(0.08)	
sex				
Gelding	Reference	Reference	Reference	
Mare	4.16(2.59)	-0.46(0.74)	-0.01(0.08)	
Stallion	9.02(4.94)	-3.09(1.42)*	0.24(0.15)	**(13.7, 6df)=3.3E-02**
clinical group				
obese$^-$lam$^-$	Reference	Reference	Reference	
obese$^-$lam$^+$	-0.93(2.95)	-0.34(0.81)	0.04(0.09)	
obese$^+$lam$^-$	-1.12(2.86)	1.84(0.81)*	-0.18(0.09)*	
obese$^+$lam$^+$	-2.72(3.37)	-0.87(0.88)	0.09(0.09)	(8.9, 9df)=4.4E-01
nutrients				
CP	-2.71(1.90)	-0.10(0.45)	-0.01(0.05)	
NDF	-0.91(2.37)	-0.05(0.49)	0.01(0.05)	
WSC	3.74(2.38)	0.04(0.48)	-0.01(0.05)	
Starch	1.16(1.81)	-0.23(0.40)	0.01(0.04)	(11.0, 12df)=5.3E-01
adiponectin	0.54(0.68)	-0.27(0.22)	0.03(0.02)	(1.6, 3df)=6.5E-01

Asterisks indicate significance of beta estimates (*<0.05, **<0.01, ***<0.001). Results obtained from a multivariate response model (responses included glucose, insulin, active GLP-1 and total GLP-1 measured at 7 time points over a 2 hour period). Time was coded in 15 minute intervals and centered at zero minutes for the trajectory estimates. Continuous predictors are scaled to zero mean and 1 standard deviation. obese$^-$lam$^-$=non-obese/no prior laminitis, obese$^-$lam$^+$=non-obese/prior laminitis, obese$^+$lam$^-$=obese/no prior laminitis, obese$^+$lam$^+$=obese/prior laminitis

Table B.9: Total GLP-1 pM trajectory multiple regression association tests

parameter	intercept estimate(se)	time estimate(se)	time^2 estimate(se)	overall signficance (wald χ^2, df) pval
mean trajectory	8.63(6.18)	3.91(0.98)	-0.41(0.10)	
sex				
Gelding	Reference	Reference	Reference	
Mare	8.72(4.97)	-1.21(0.83)	0.10(0.08)	
Stallion	15.81(9.65)	-5.11(1.62)**	0.47(0.17)**	**(14.9, 6df)=2.1E-02**
clinical group				
obese$^-$lam$^-$	Reference	Reference	Reference	
obese$^-$lam$^+$	-8.06(5.59)	0.15(0.91)	0.01(0.09)	
obese$^+$lam$^-$	-5.36(5.46)	2.45(0.90)**	-0.23(0.09)*	
obese$^+$lam$^+$	-6.33(6.38)	-1.54(1.01)	0.17(0.10)	**(16.9, 9df)=5.0E-02**
nutrients				
CP	-5.83(3.48)	-0.06(0.51)	-0.06(0.05)	
NDF	-0.82(4.21)	0.63(0.55)	-0.06(0.06)	
WSC	1.28(4.22)	0.33(0.55)	-0.01(0.06)	
Starch	4.17(3.23)	-0.58(0.45)	0.00(0.05)	**(35.5, 12df)=3.9E-04**
NEFA	0.45(1.51)	0.00(0.30)	-0.01(0.03)	(1.5, 3df)=6.8E-01
triglycerides	0.07(1.43)	-0.53(0.28)	0.05(0.03)	(5.3, 3df)=1.5E-01
GCG haplotype				
haploptype A	Reference	Reference	Reference	
haploptype B	4.63(2.16)*	-0.96(0.42)*	0.09(0.05)*	
haploptype C	1.34(2.05)	-0.33(0.40)	0.02(0.04)	
haploptype D	-2.18(3.82)	0.01(0.75)	0.00(0.08)	
haploptype E	5.81(3.10)	-1.06(0.61)	0.03(0.07)	**(22.2, 12df)=3.6E-02**

Asterisks indicate significance of beta estimates (*<0.05, **<0.01, ***<0.001). Results obtained from a multivariate response model (responses included glucose, insulin, active GLP-1 and total GLP-1 measured at 7 time points over a 2 hour period). Time was coded in 15 minute intervals and centered at zero minutes for the trajectory estimates. Continuous predictors are scaled to zero mean and 1 standard deviation. obese$^-$lam$^-$=non-obese/no prior laminitis, obese$^-$lam$^+$=non-obese/prior laminitis, obese$^+$lam$^-$=obese/no prior laminitis, obese$^+$lam$^+$=obese/prior laminitis

Appendix C

Supplemental Material: Chapter 4 Improved linear mixed model of polygenic traits in populations with familial relatedness

Figure C.1: Enlarged version of Figure 4.2b Morgan horse PCA plot with numeric representation of different farms.

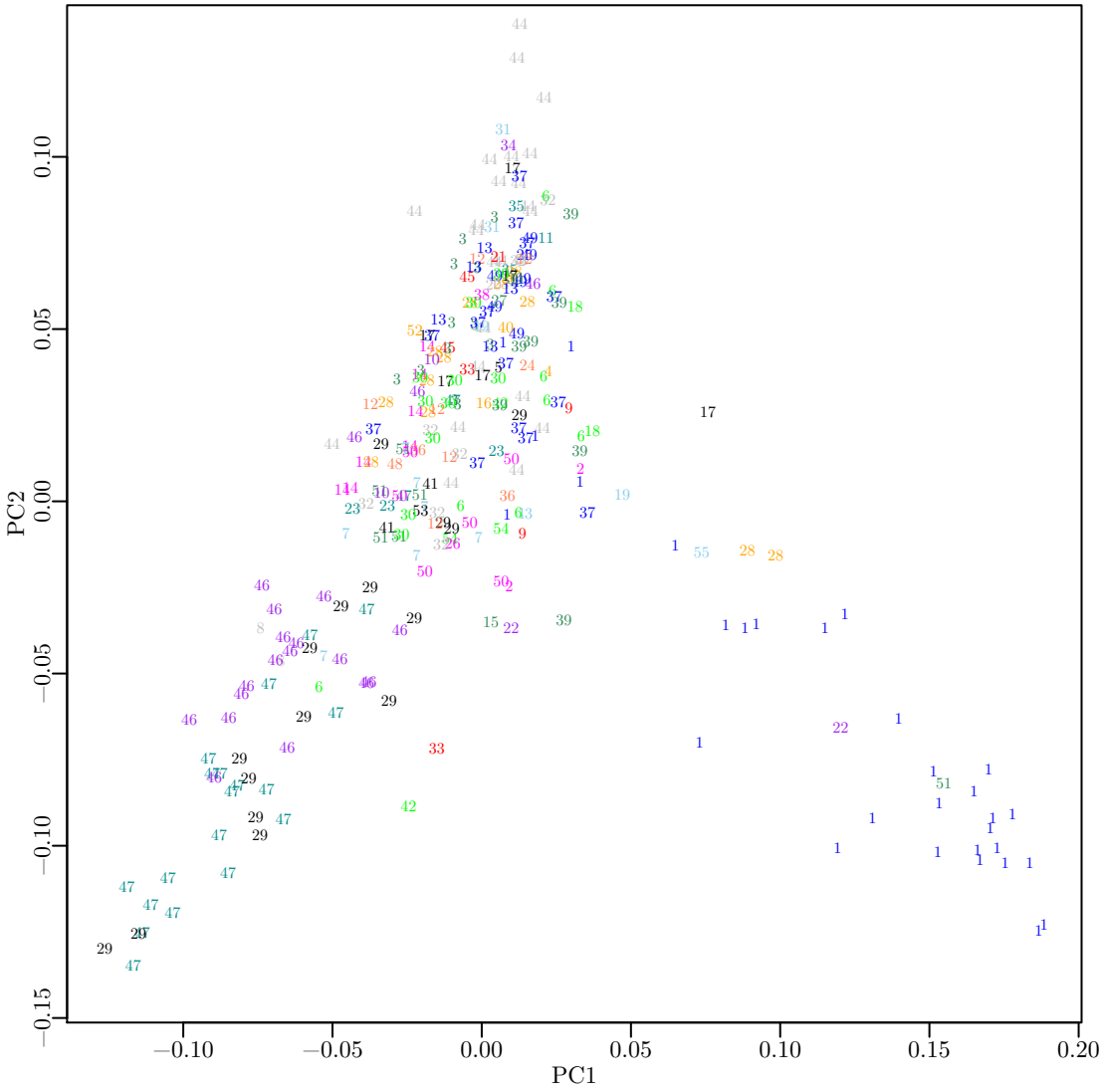


Figure C.2: Enlarged version of Figure 4.2c Welsh pony PCA plot with numeric representation of different farms.

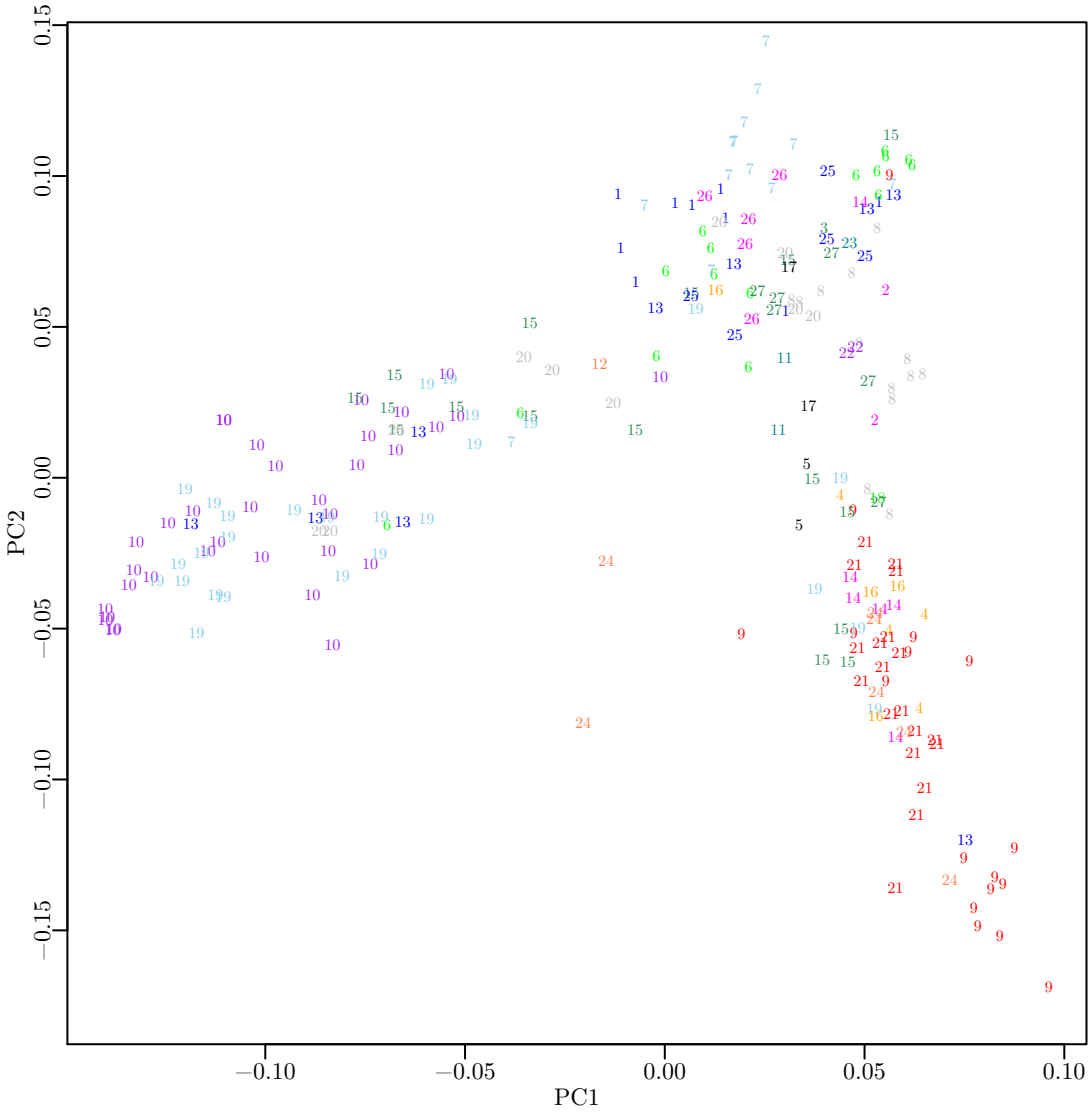


Figure C.3: Manhattan plot displaying the standard bayesian sparse linear mixed model results for the simulated QTL-MAS dataset. Red vertical lines indicate the location of true QTL.

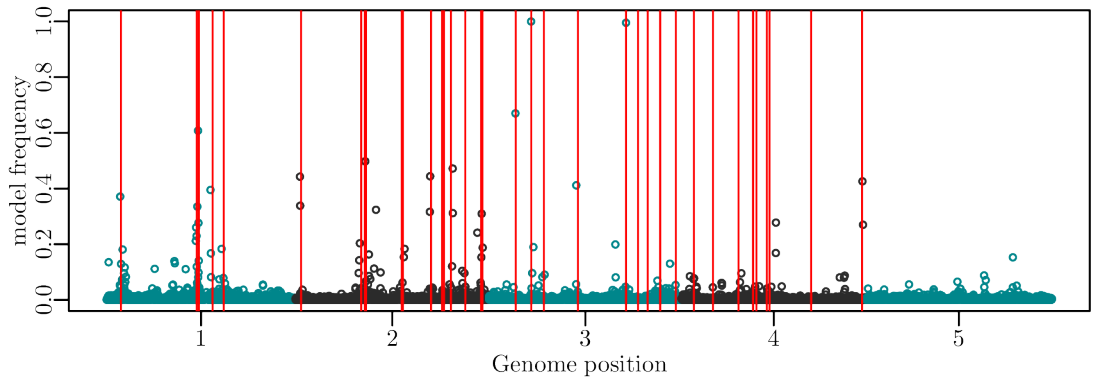


Figure C.4: Manhattan plot displaying the improved linear mixed model results for the simulated QTL-MAS dataset using a 200 kilobase bin size. Red vertical lines indicate the location of true QTL. Genome-wide p-value **<0.05** (solid grey horizontal line) and suggestive (dashed grey horizontal line) thresholds are also shown.

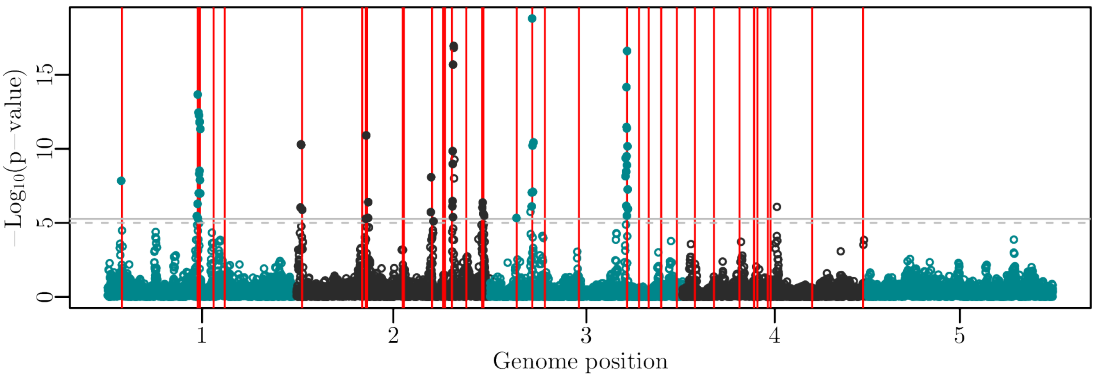
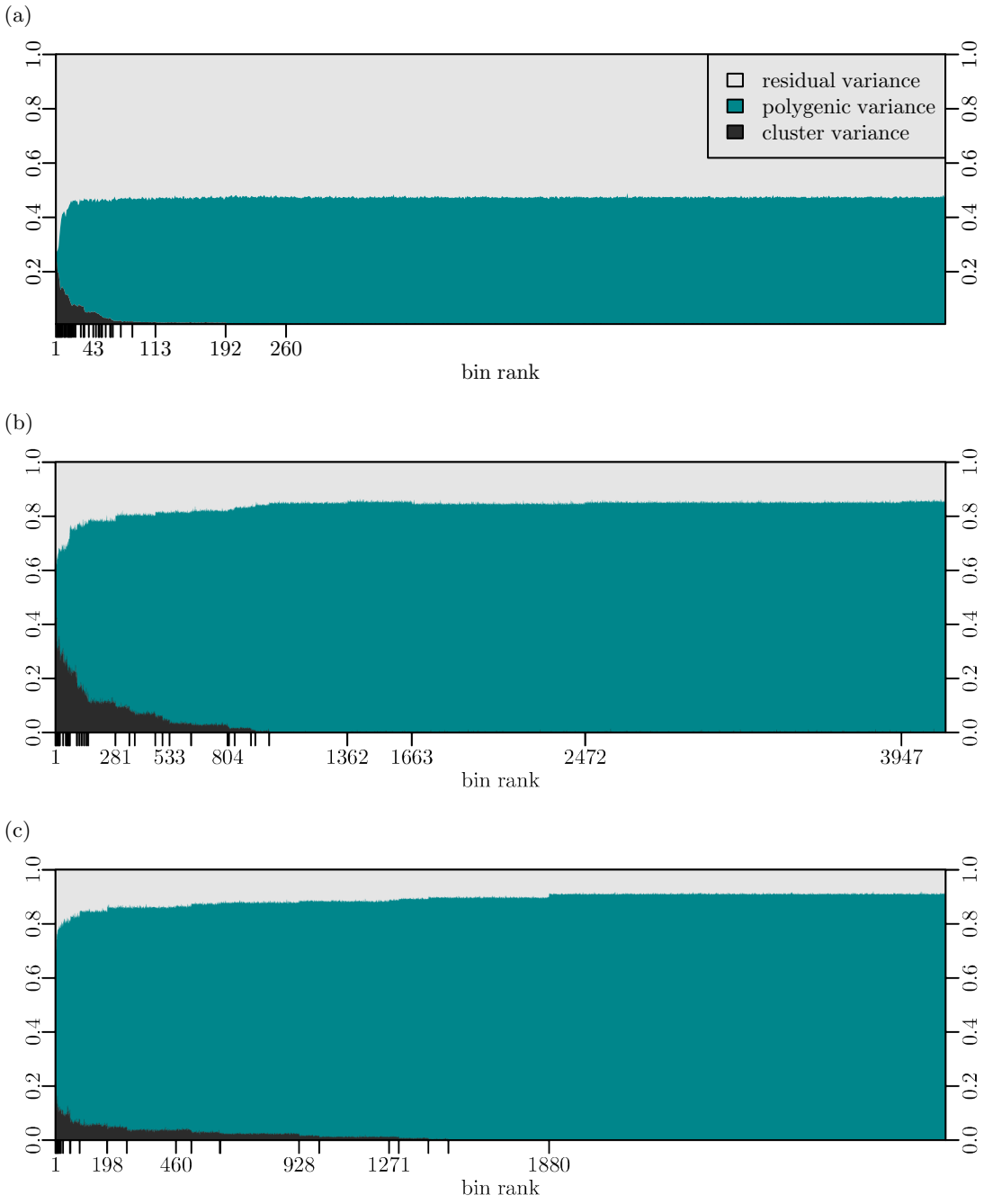


Figure C.5: Illustration of variance decomposition throughout the stepwise feature SNP selection process for the a) QTL-MAS simulated dataset, b) Morgan horse real height dataset, and c) Welsh pony real height dataset. Tick marks along the X-axis represent selected bins.



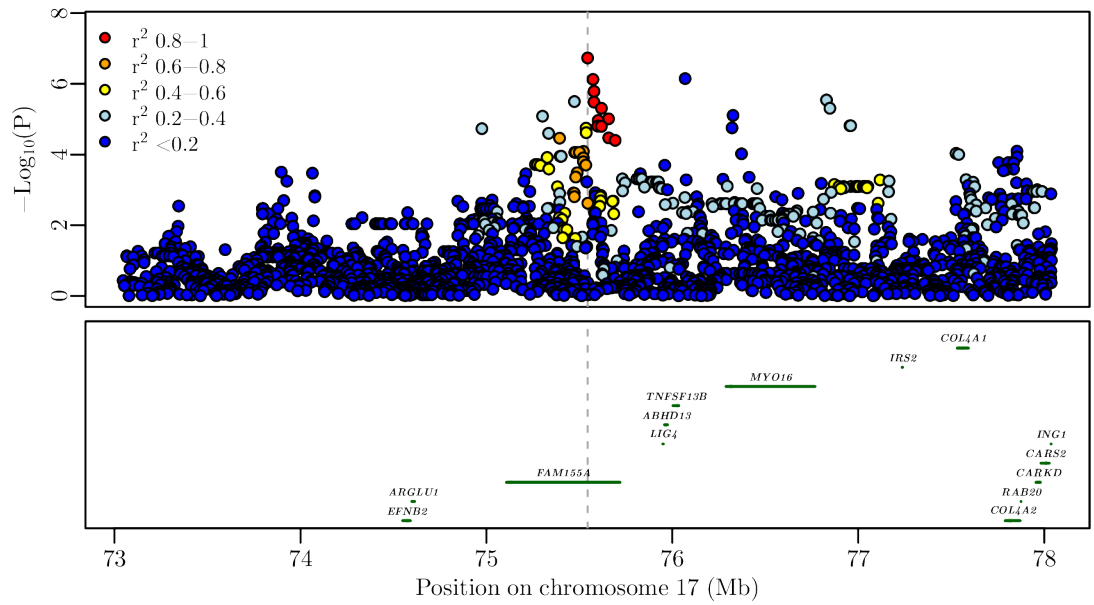
Appendix D

Supplemental Material: Chapter 5 Genome wide association study of equine metabolic trait variation

D.1 Regional plots of EMS candidate loci. Color-coding represents correlation (r^2 value) with the reference SNP (dashed line)

Figure D.1: Neck circumference:height ratio loci

(a)



(b)

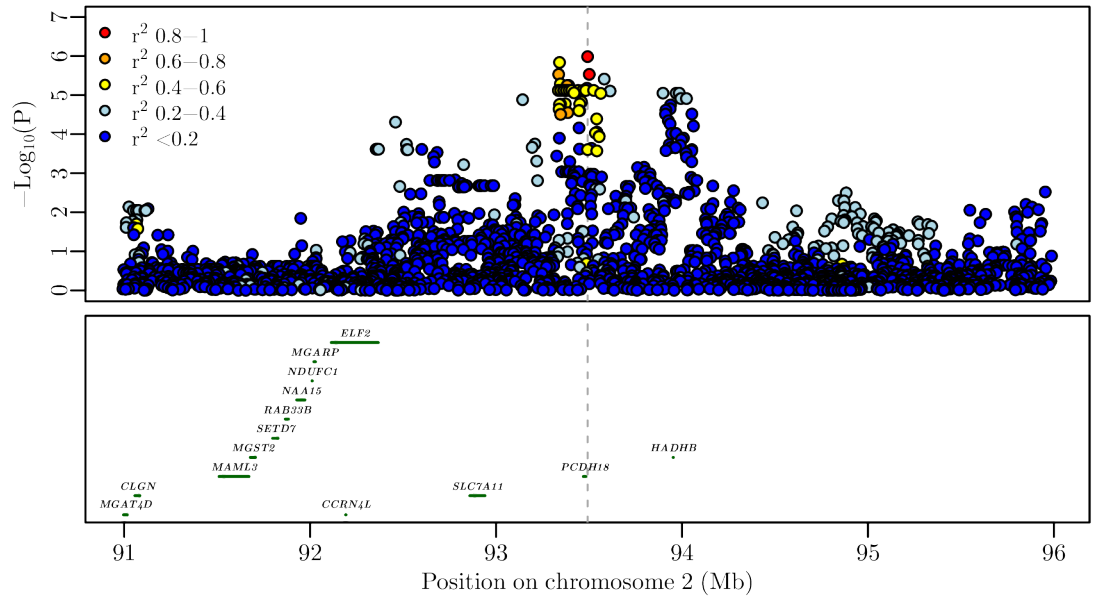
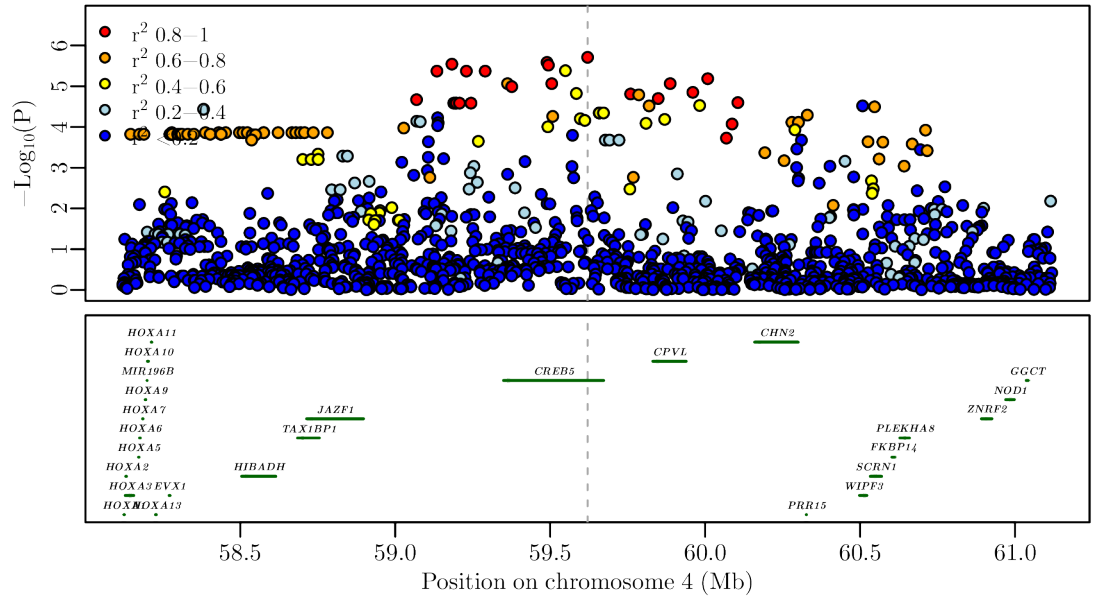


Figure D.1 *Continued from previous page*: Neck circumference:height ratio loci

(c)



(d)

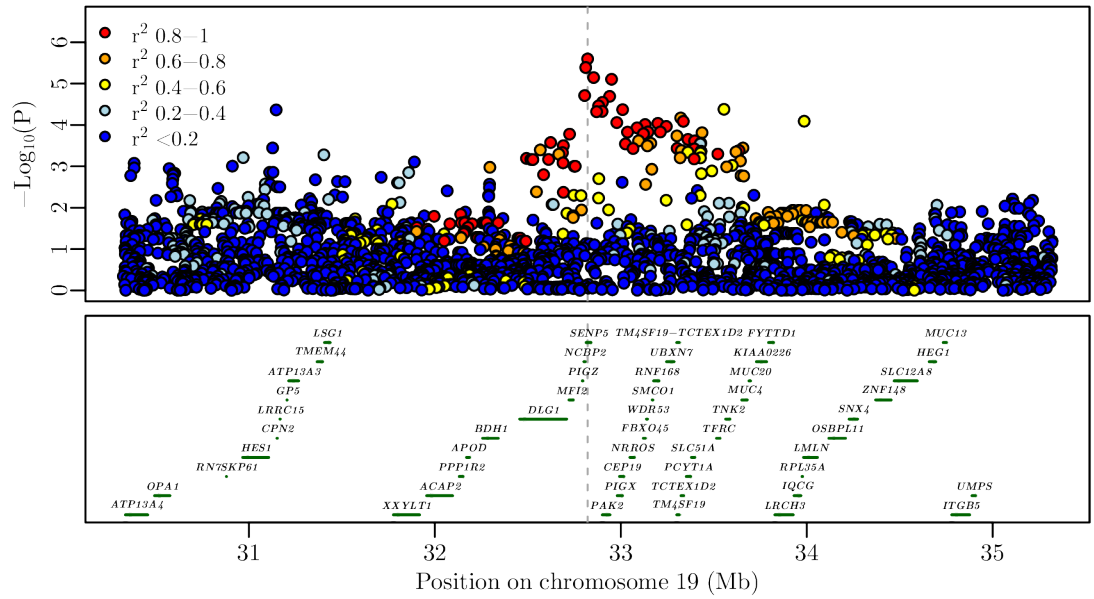
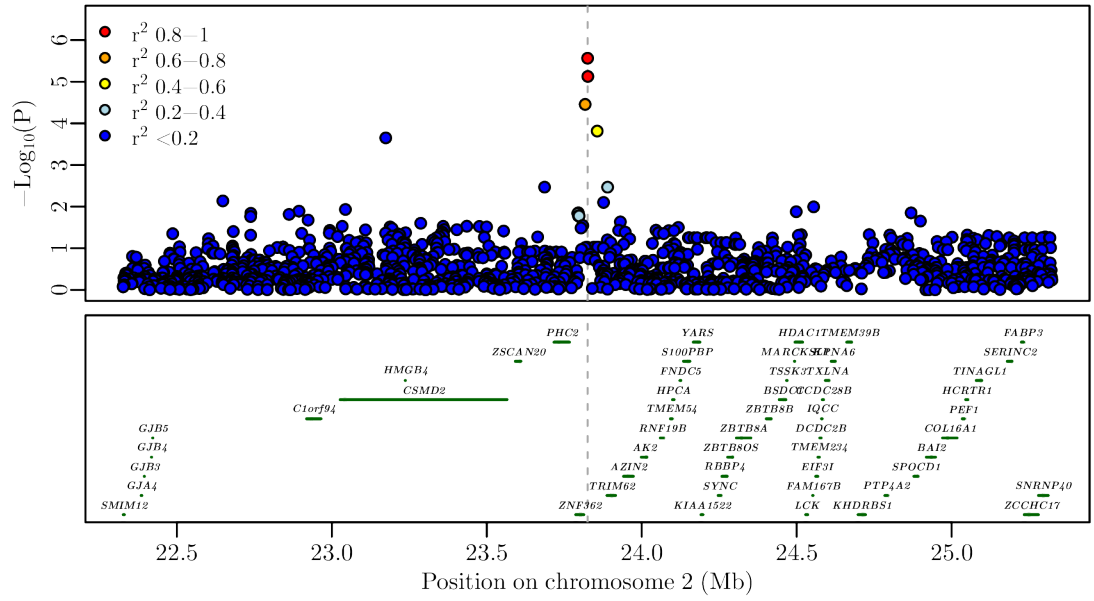


Figure D.1 *Continued from previous page*: Neck circumference:height ratio loci

(e)



(f)

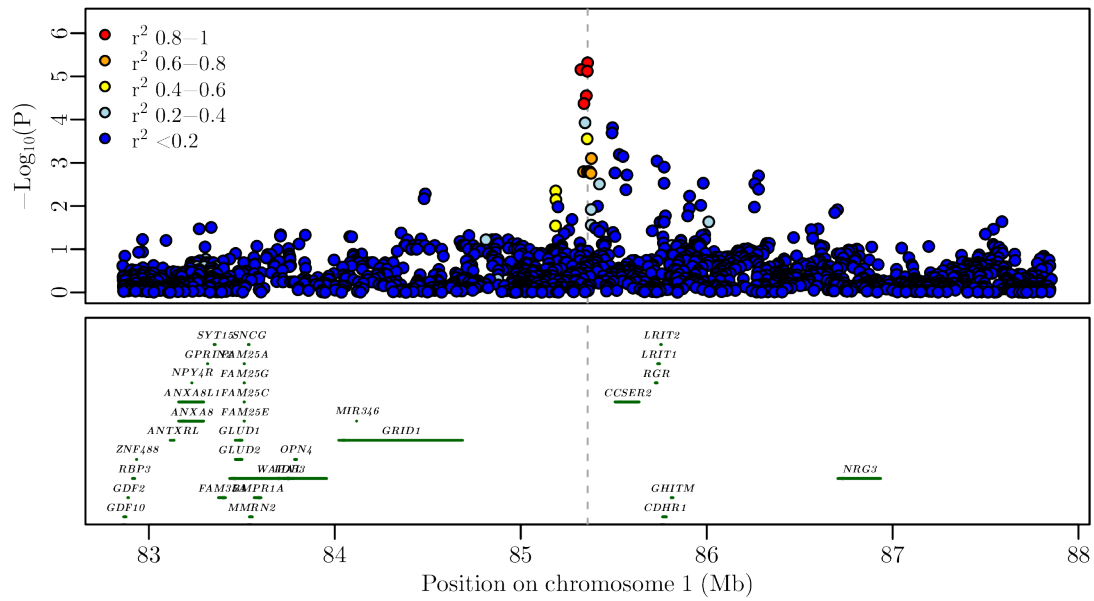
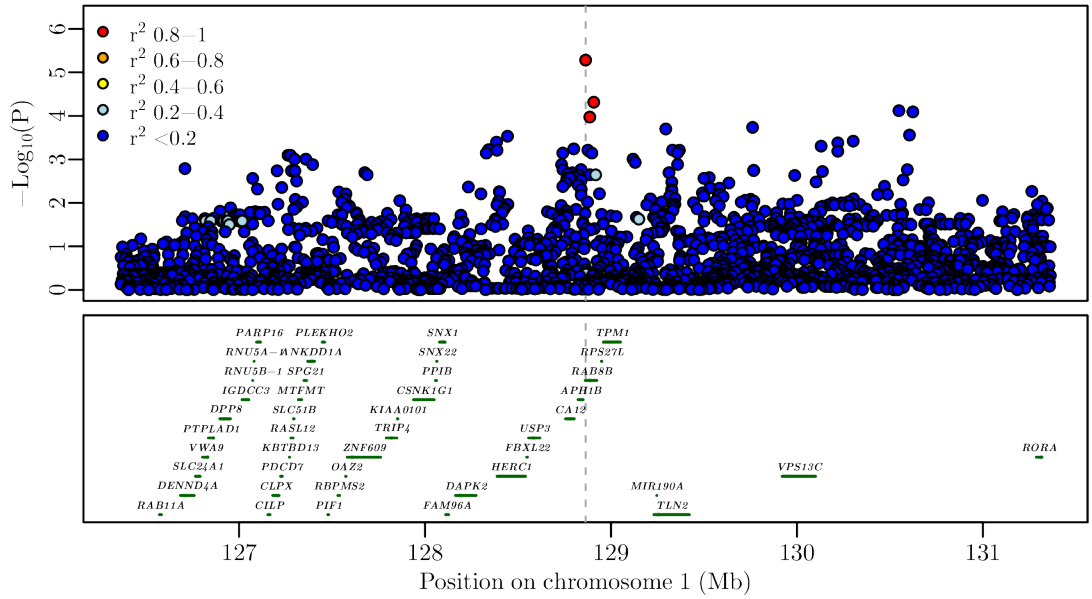


Figure D.1 *Continued from previous page*: Neck circumference:height ratio loci

(g)



(h)

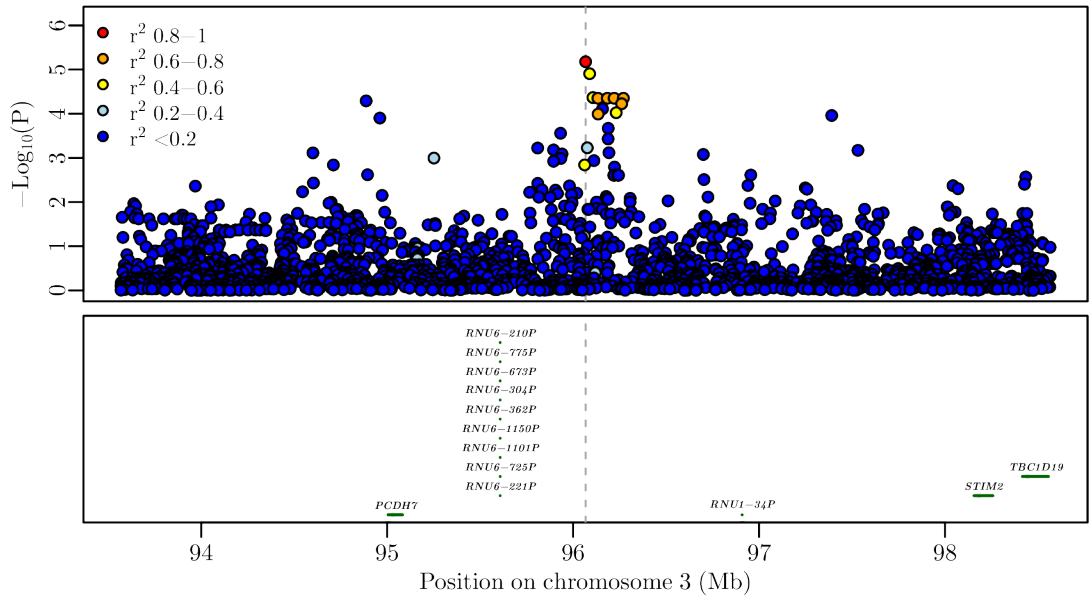


Figure D.1 *Continued from previous page*: Neck circumference:height ratio loci

(i)

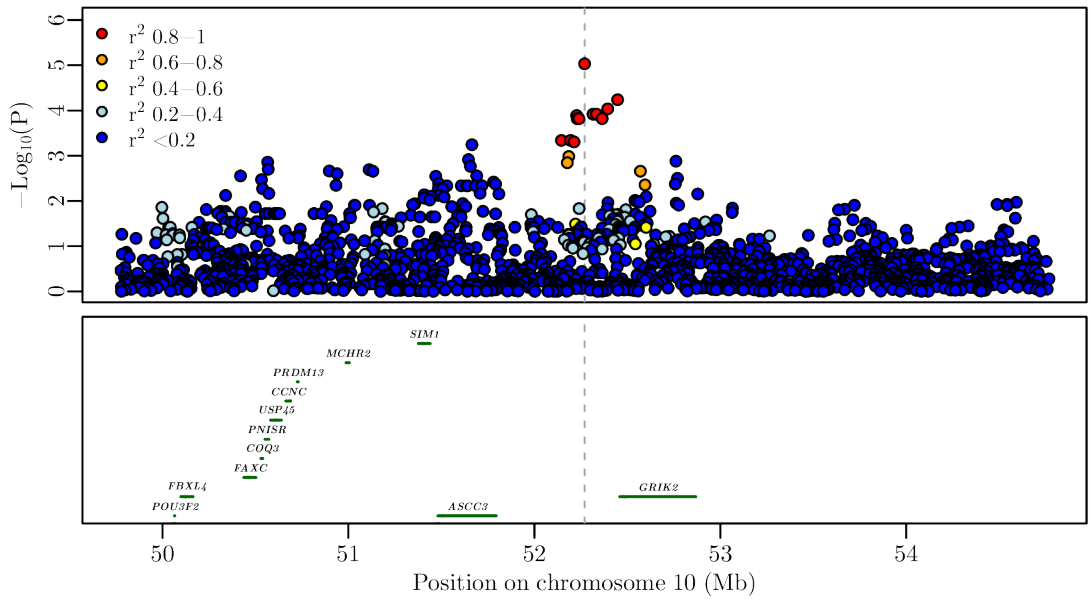
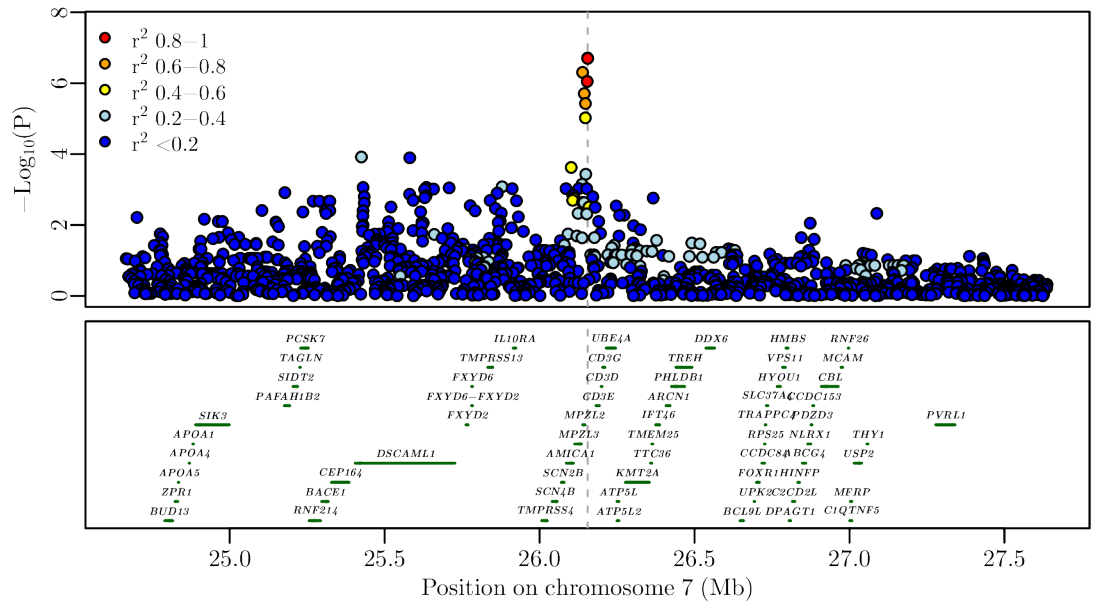


Figure D.2: Girth:height ratio loci

(a)



(b)

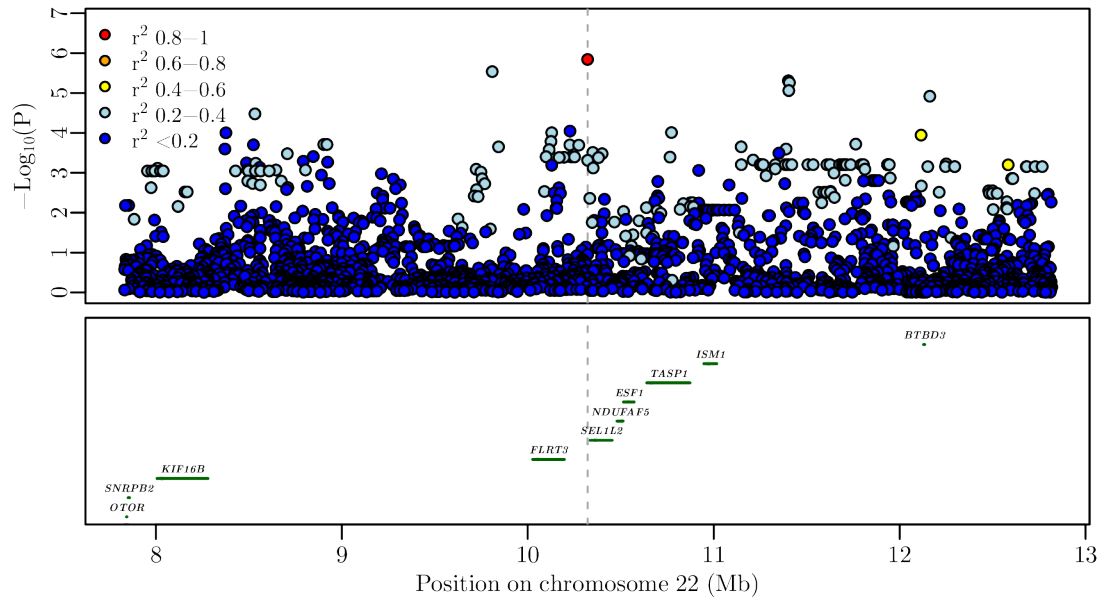
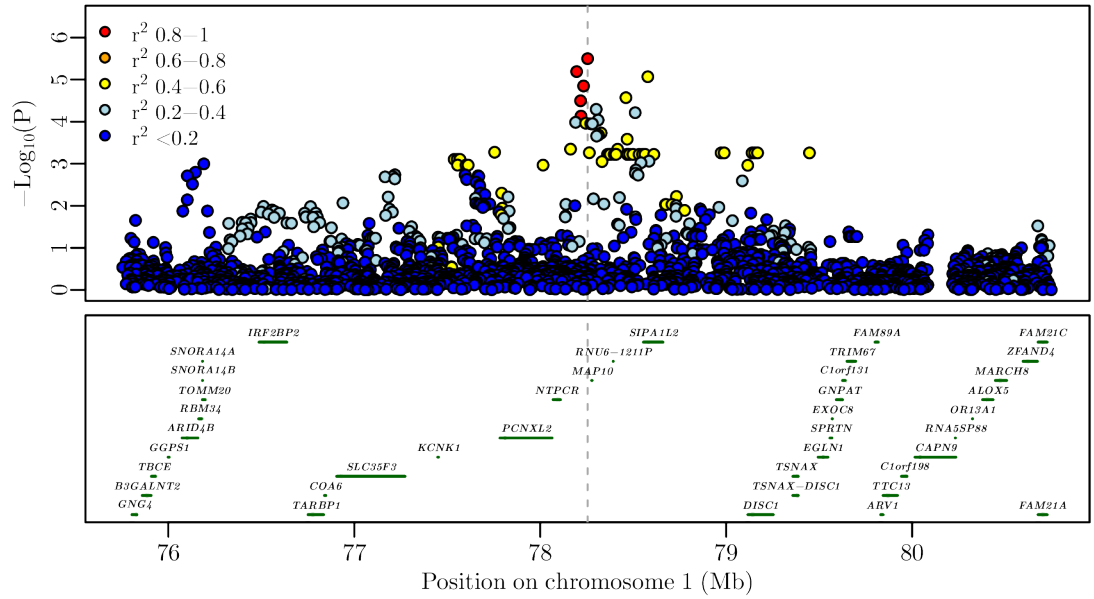


Figure D.2 *Continued from previous page*: Girth:height ratio loci

(c)



(d)

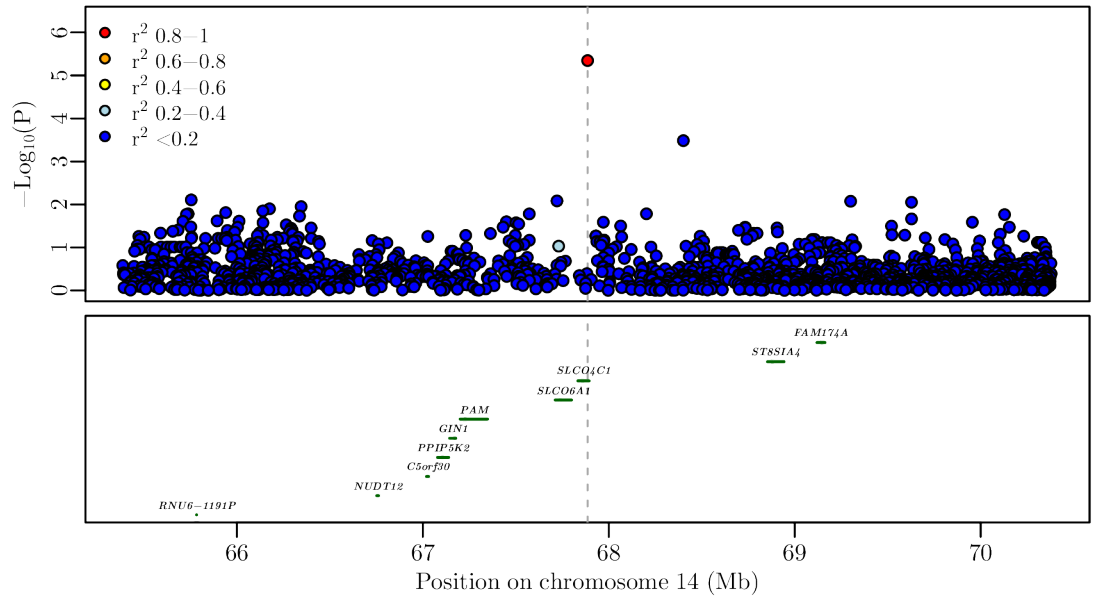
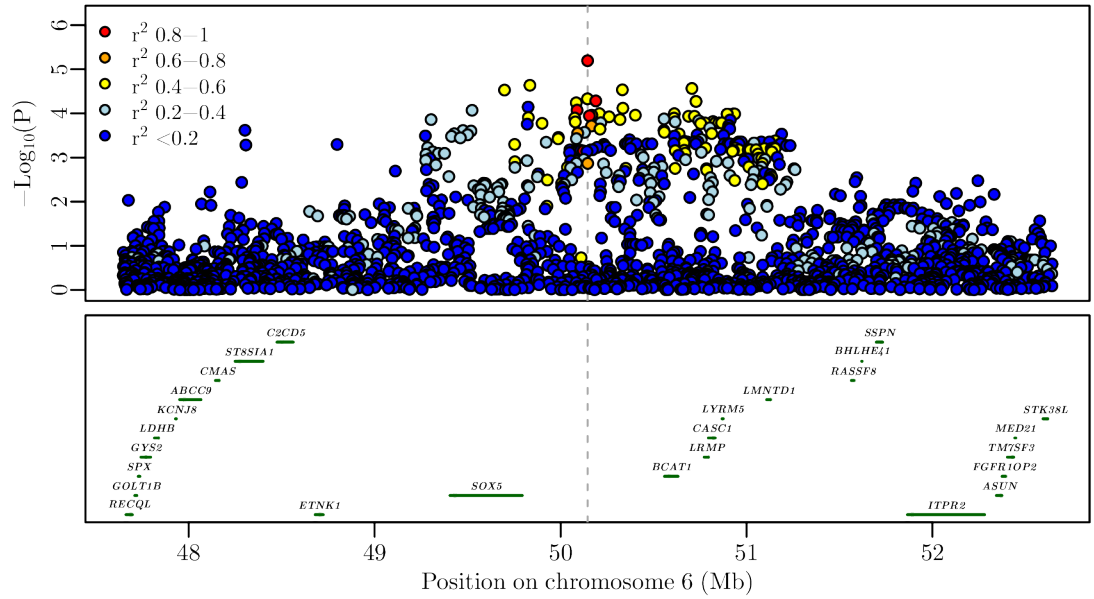


Figure D.2 *Continued from previous page*: Girth:height ratio loci

(e)



(f)

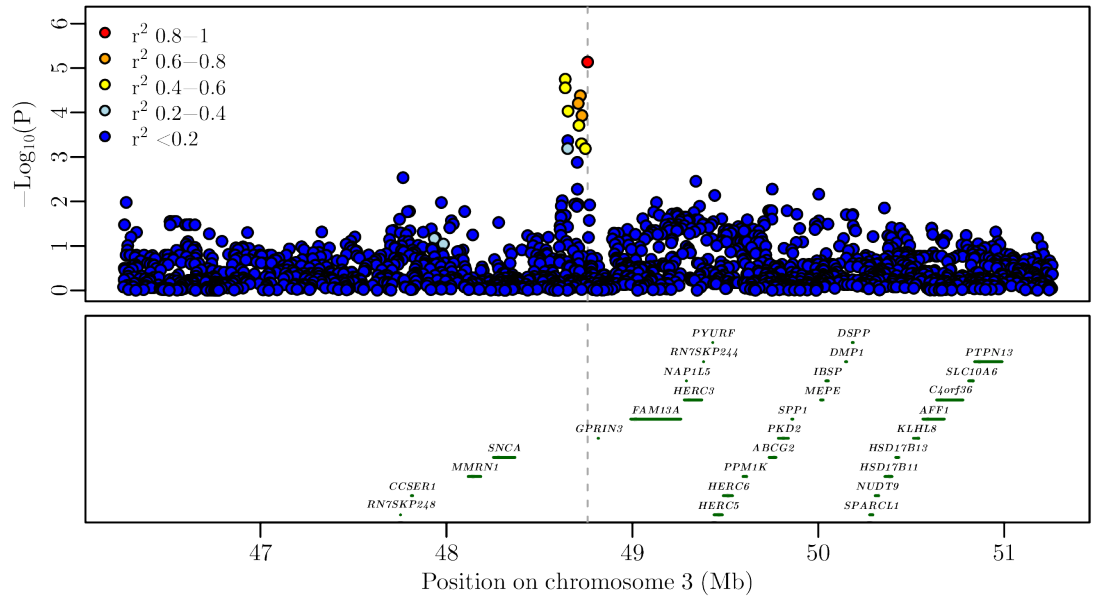
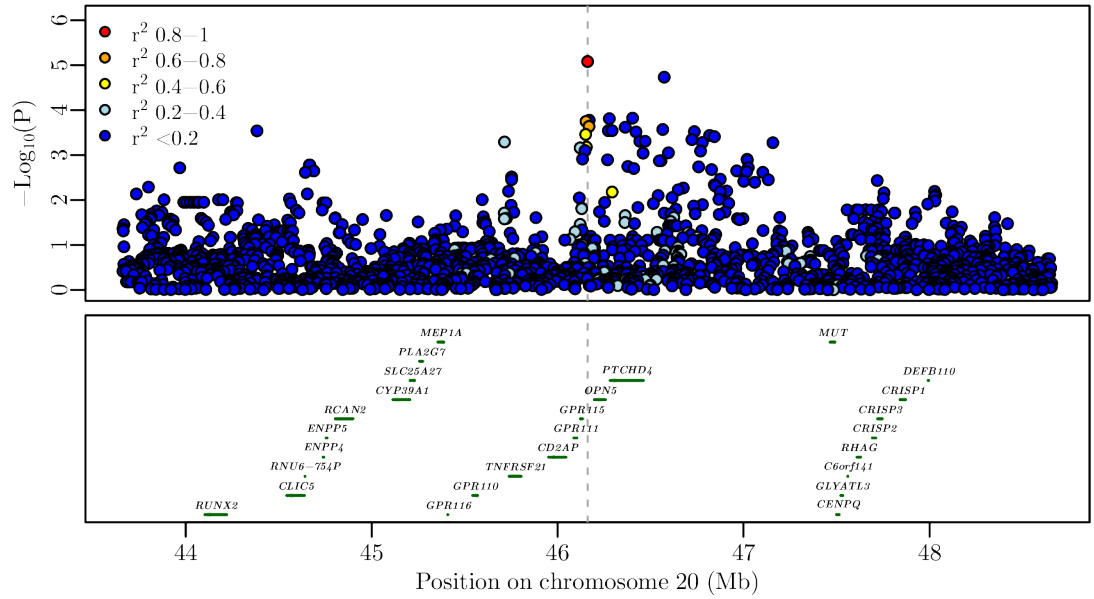


Figure D.2 *Continued from previous page*: Girth:height ratio loci

(g)



(h)

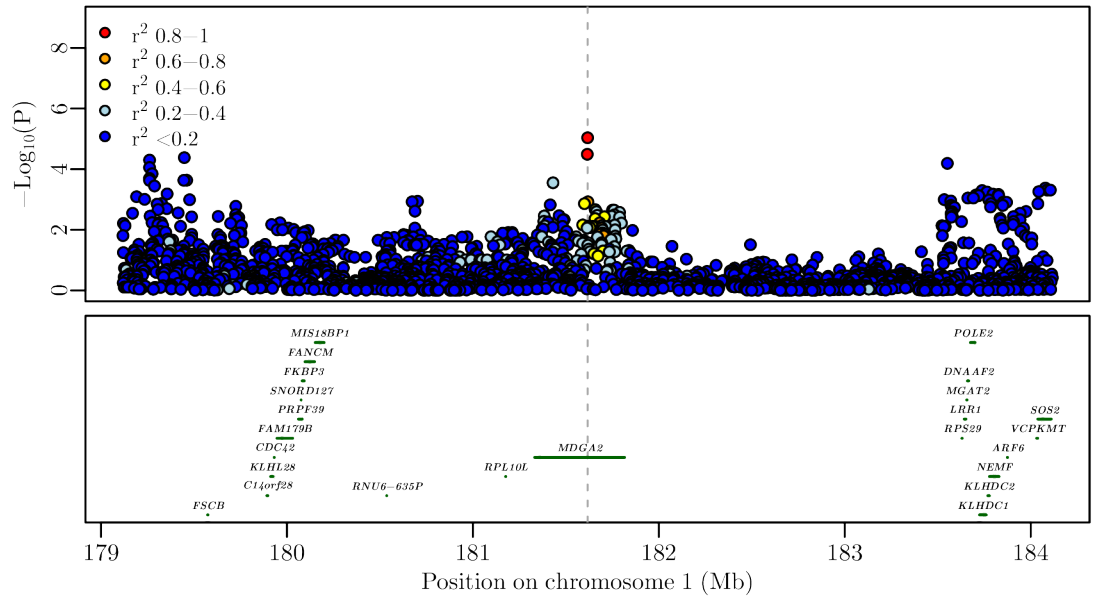


Figure D.3: Fasting glucose loci

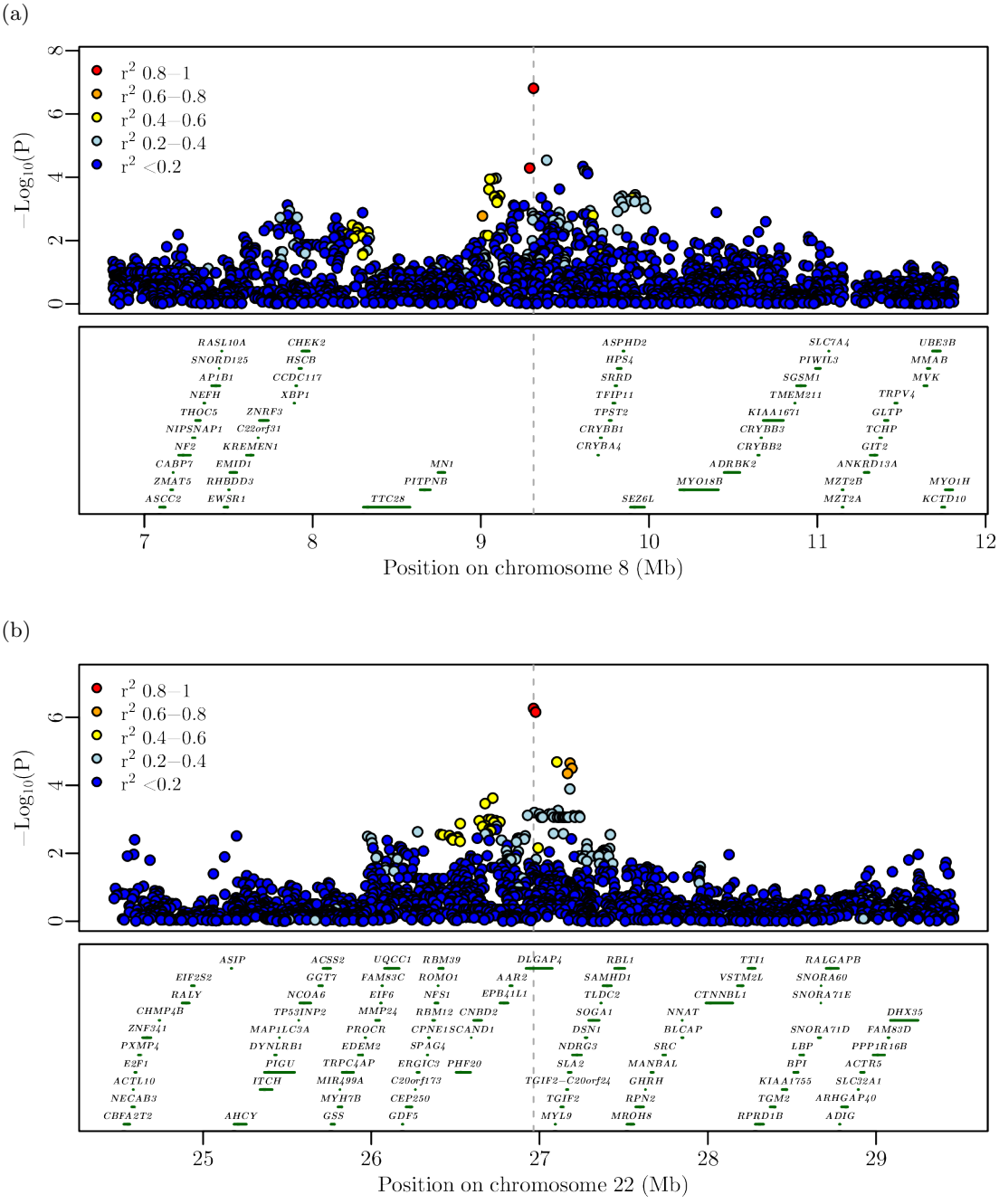
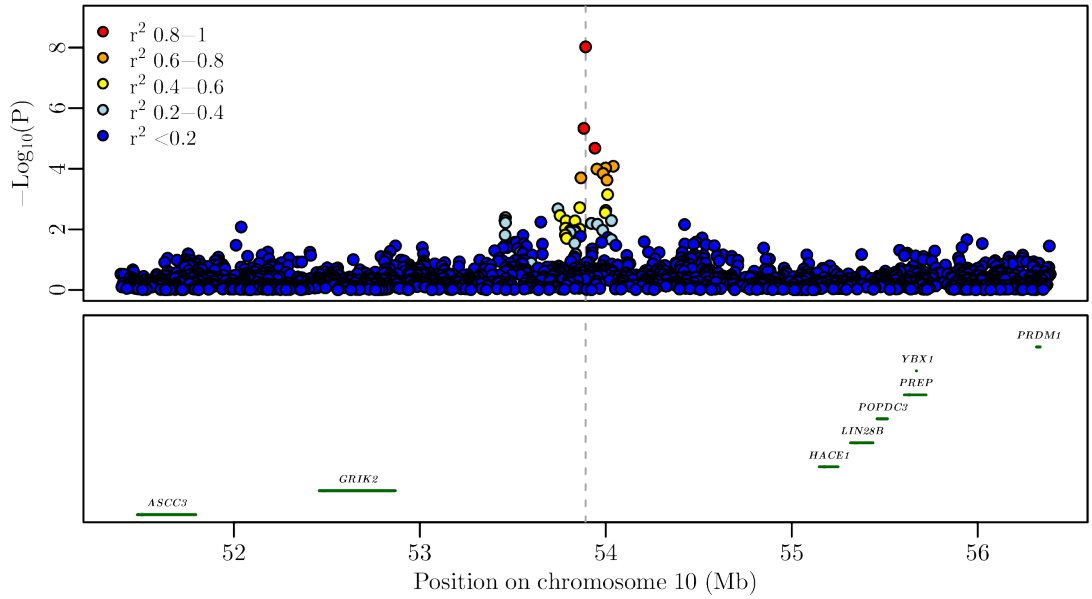


Figure D.4: Fasting insulin loci

(a)



(b)

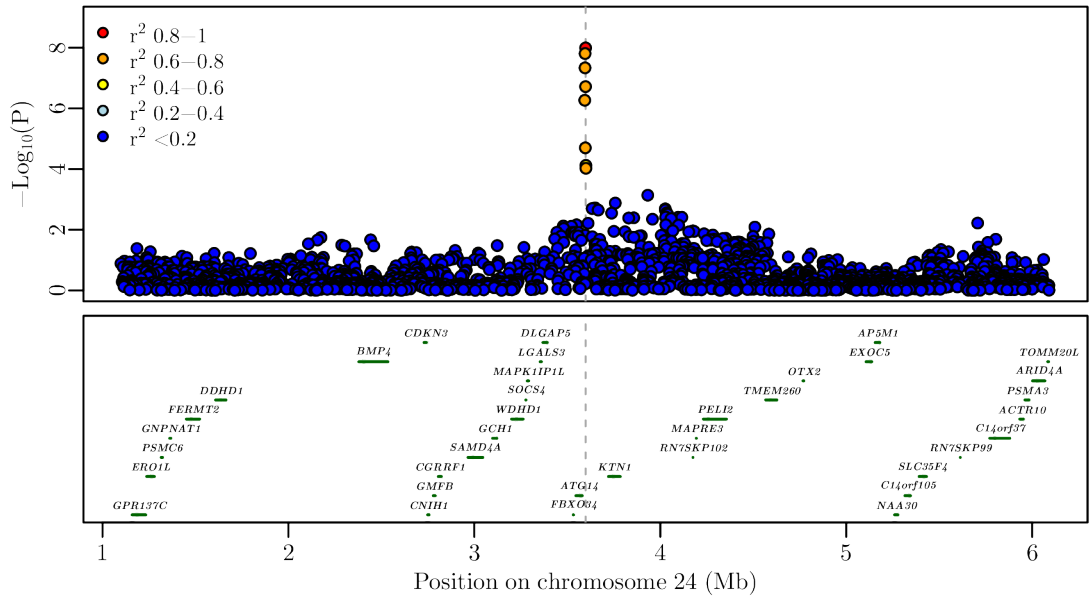
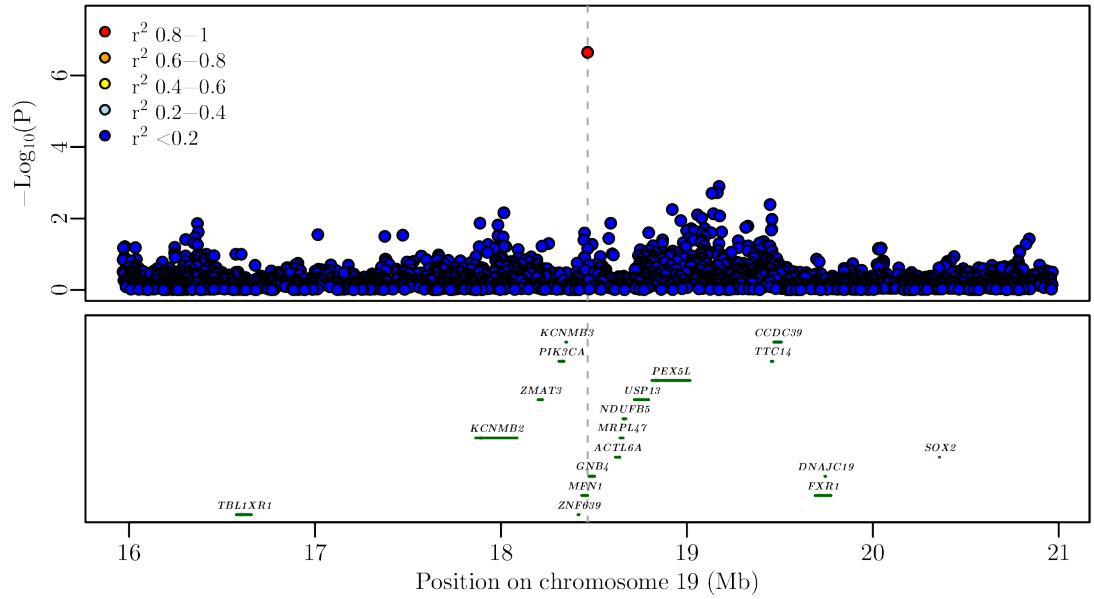


Figure D.4 *Continued from previous page*: Fasting insulin loci

(c)



(d)

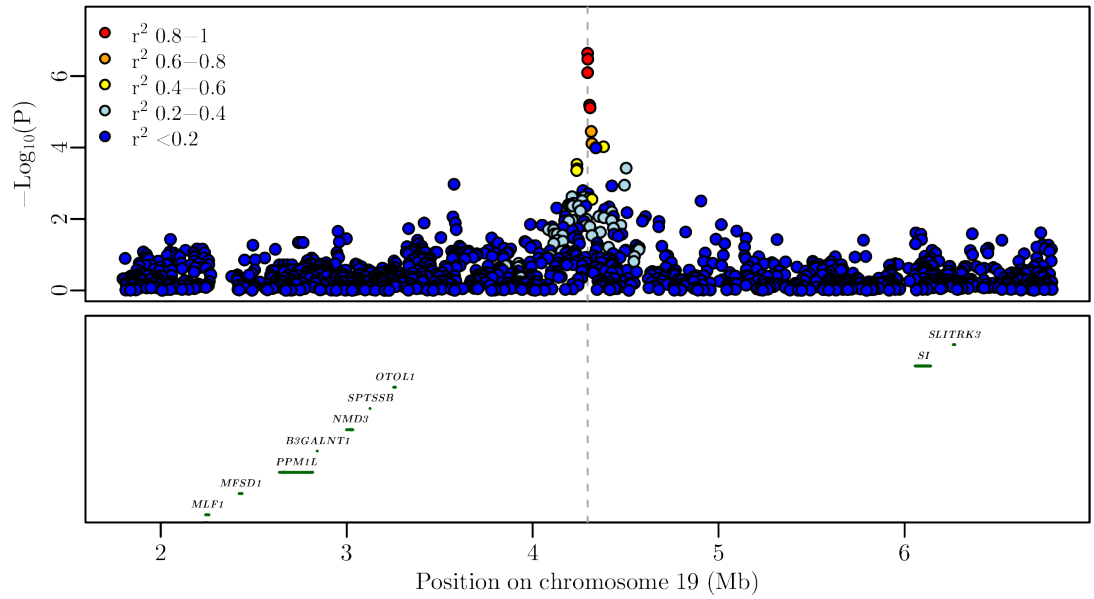
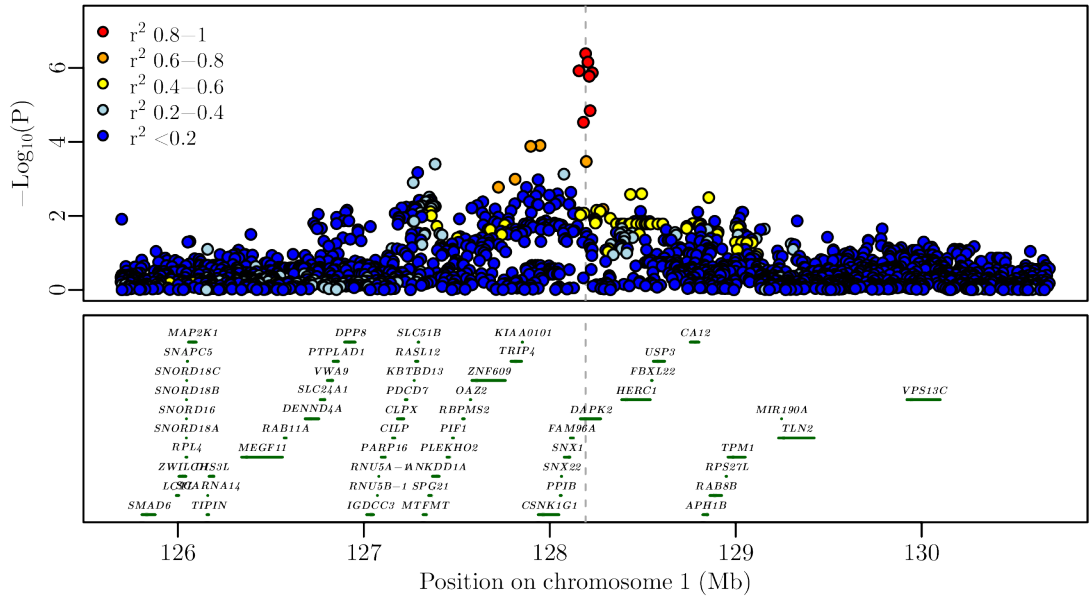


Figure D.4 *Continued from previous page*: Fasting insulin loci

(e)



(f)

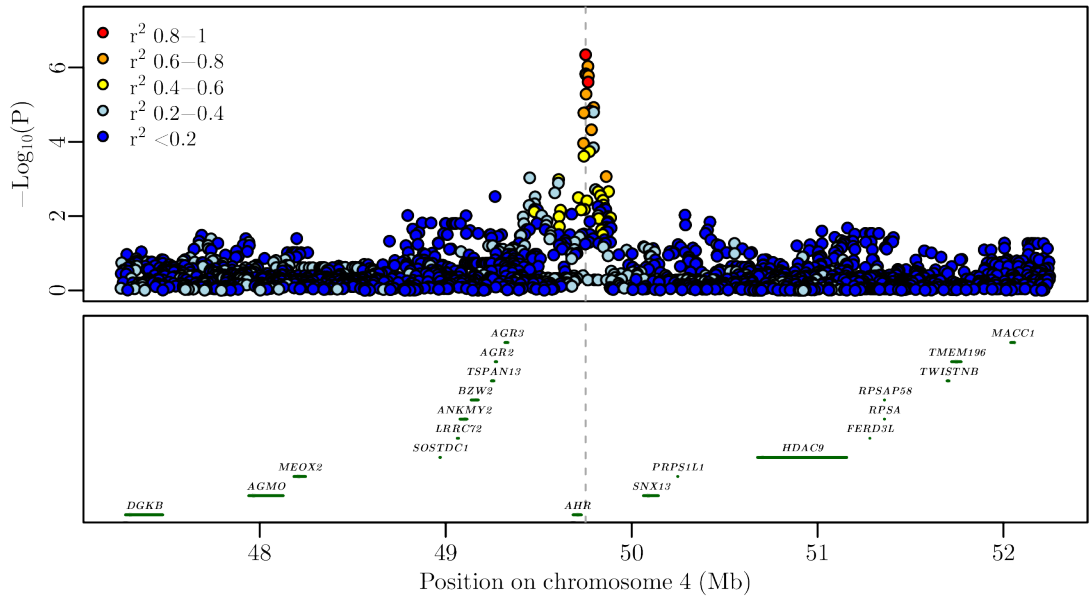
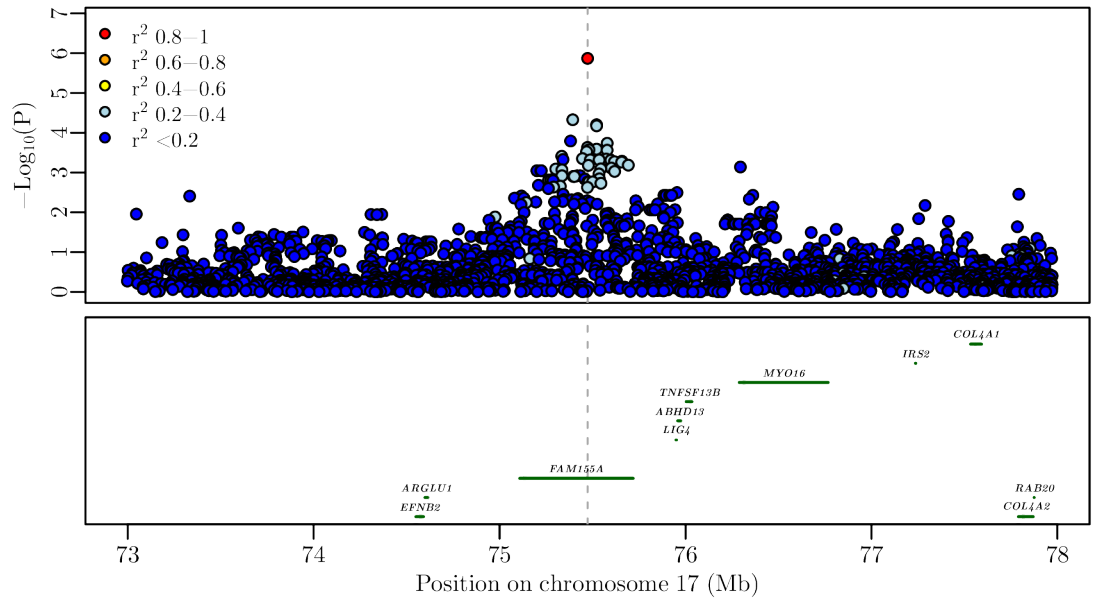


Figure D.4 *Continued from previous page*: Fasting insulin loci

(g)



(h)

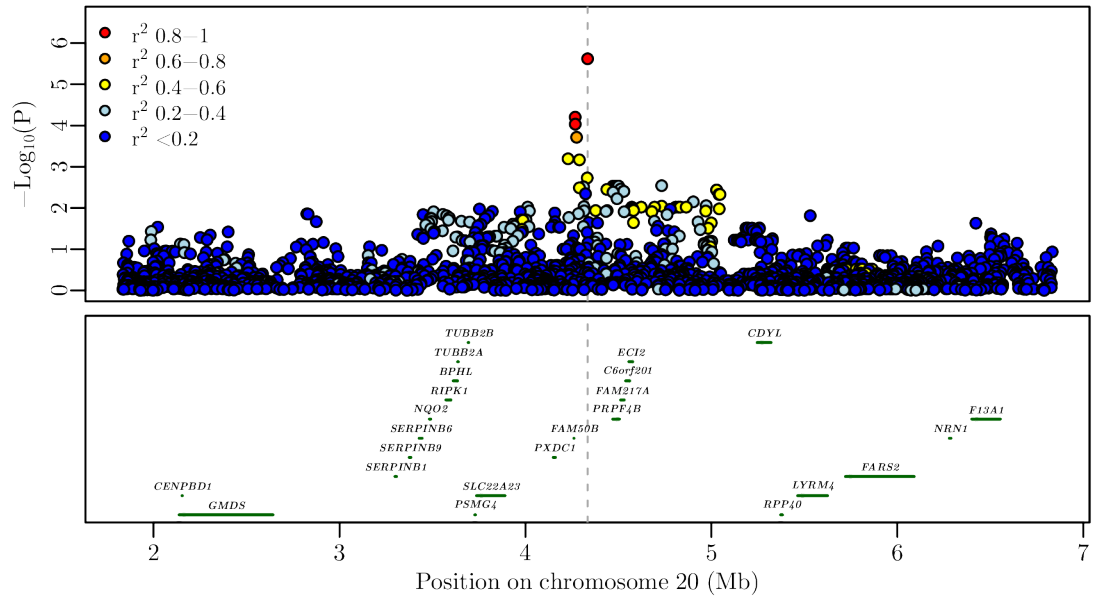
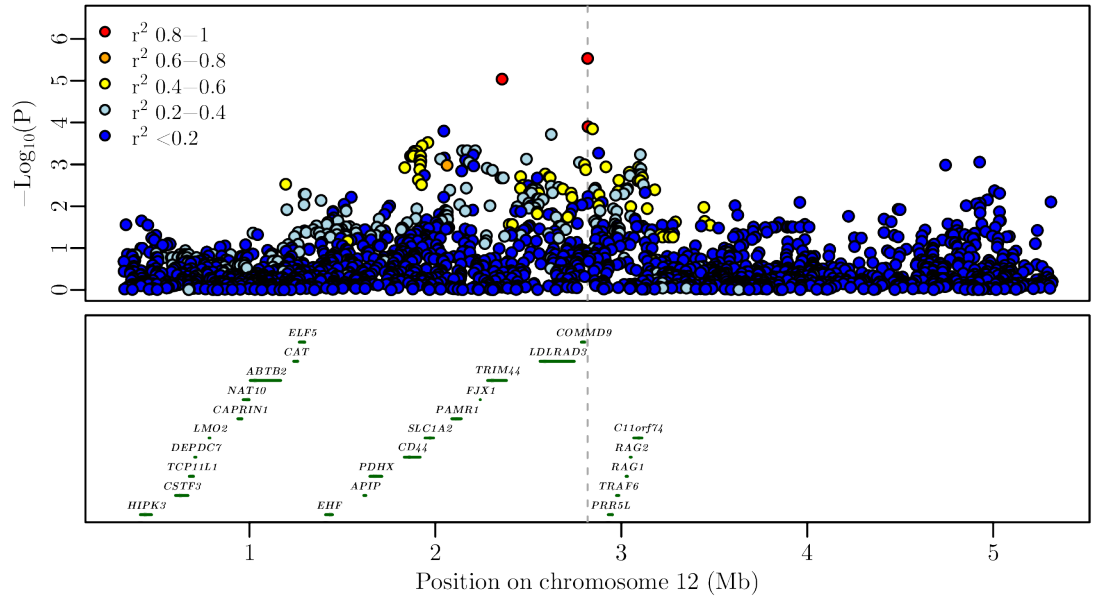


Figure D.4 *Continued from previous page*: Fasting insulin loci

(i)



(j)

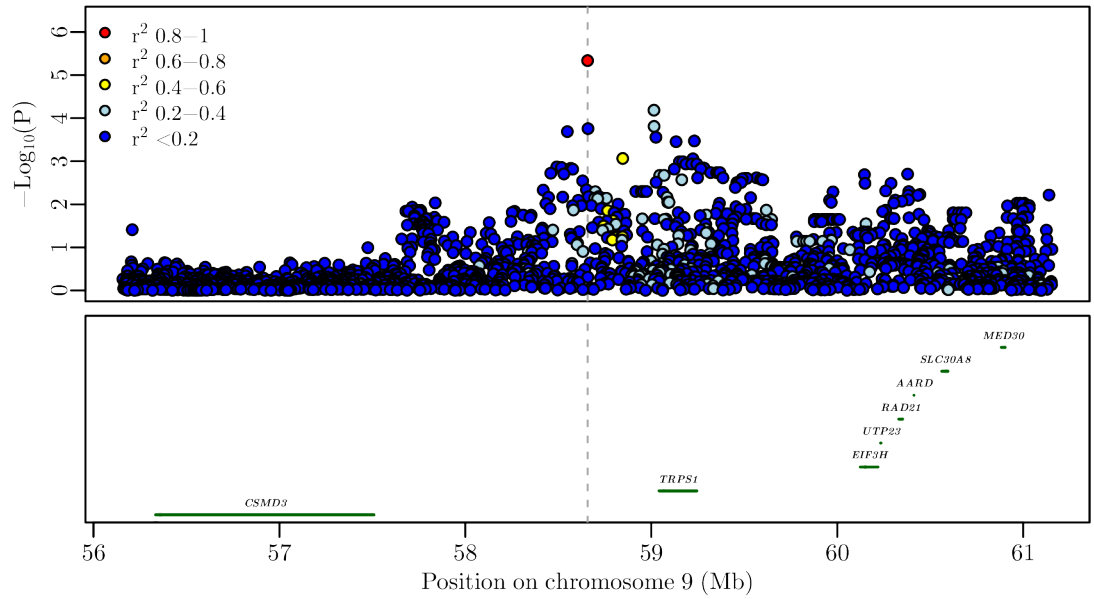
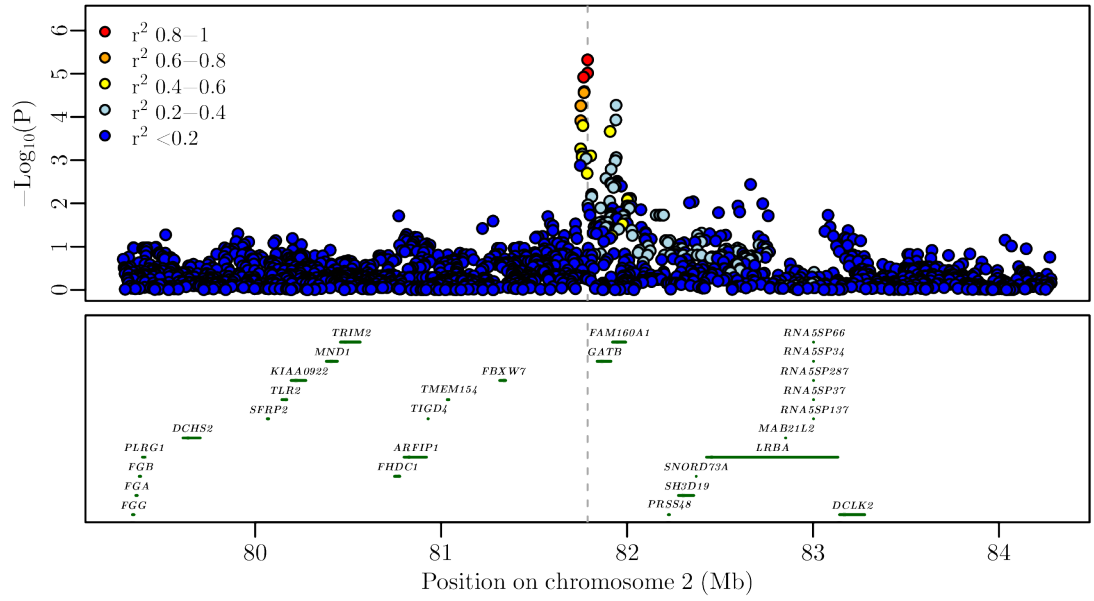


Figure D.4 *Continued from previous page*: Fasting insulin loci

(k)



(l)

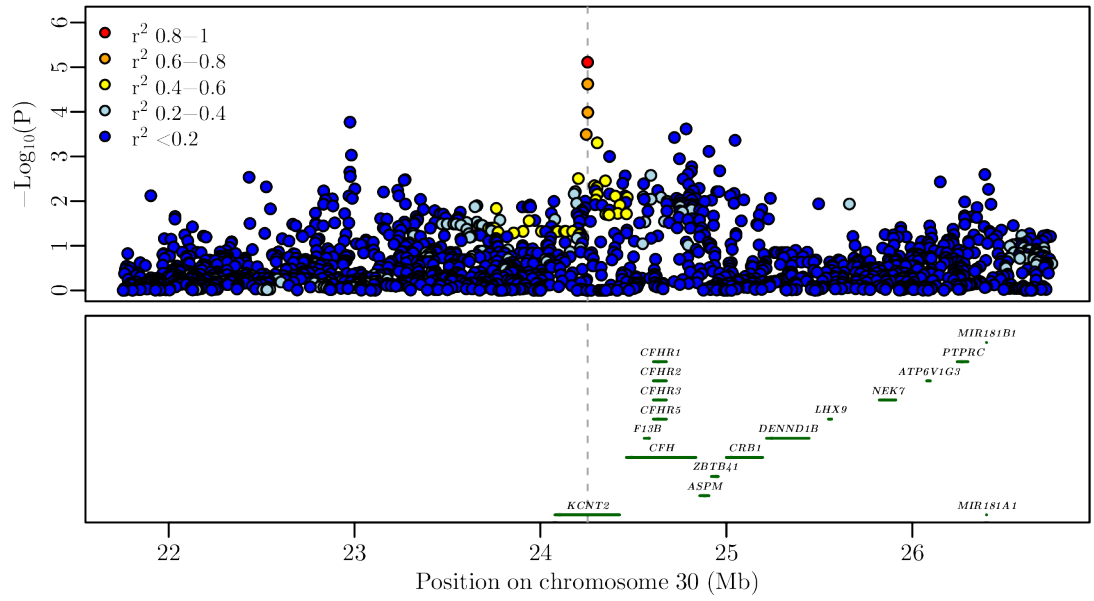
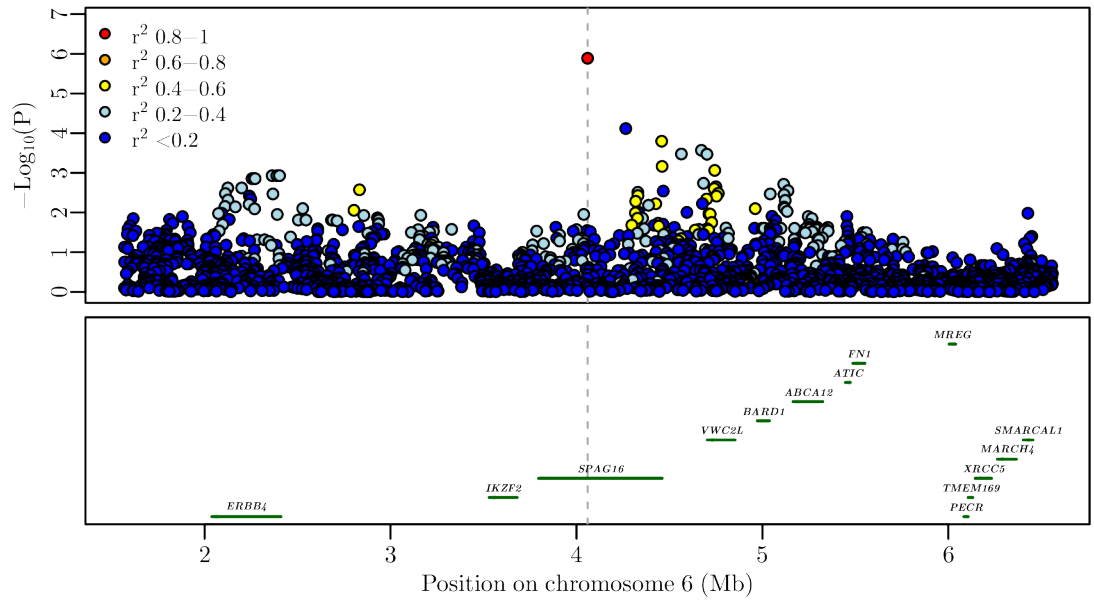


Figure D.5: Glucose 75-minute post-oral sugar test loci

(a)



(b)

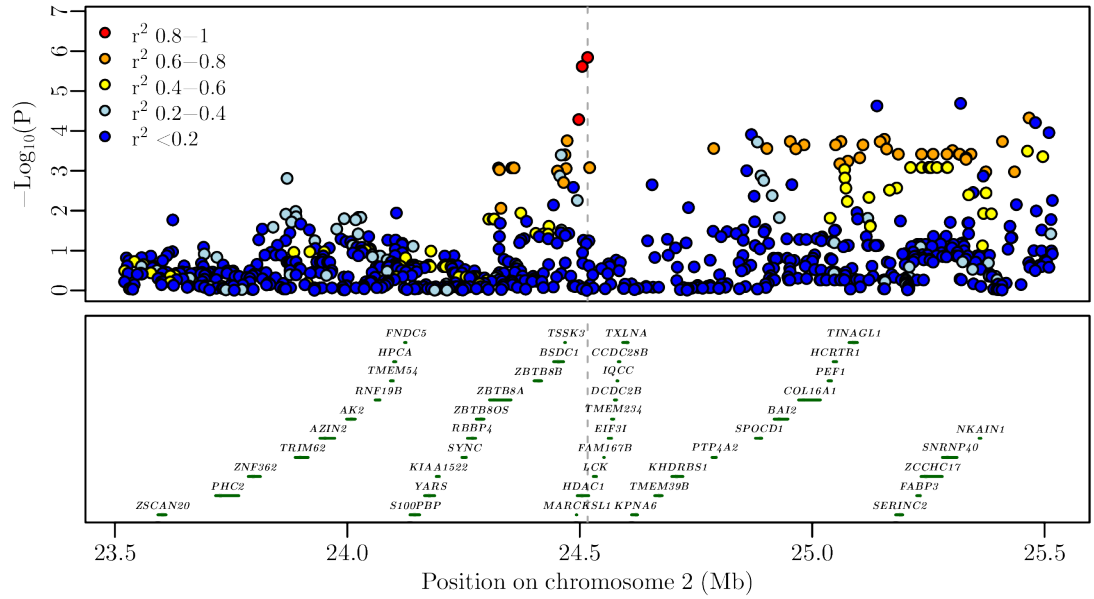
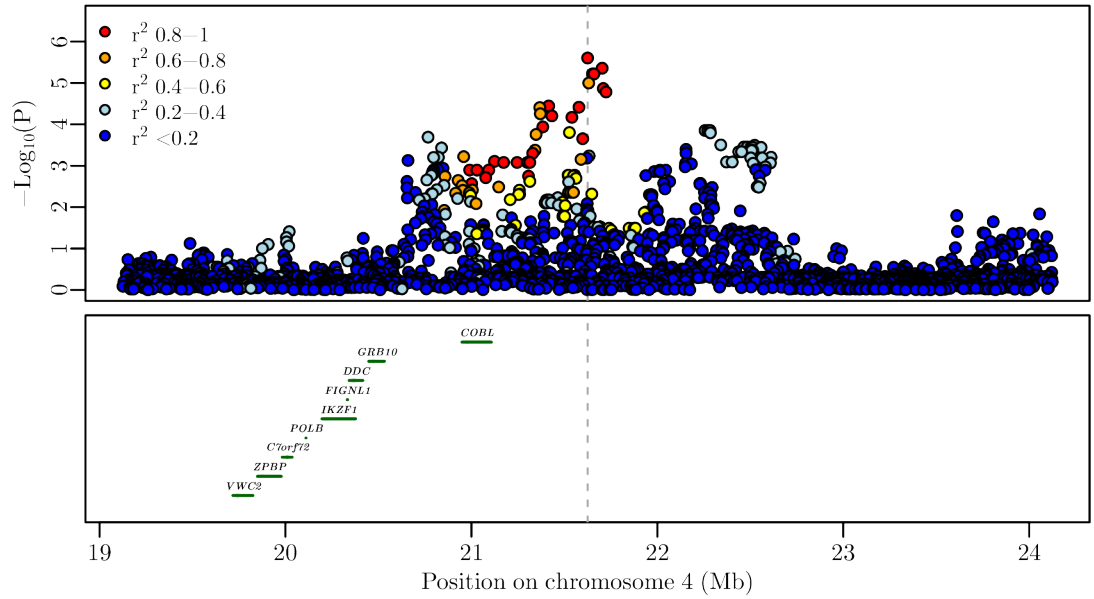


Figure D.5 *Continued from previous page*: Glucose 75-minute post-oral sugar test loci

(c)



(d)

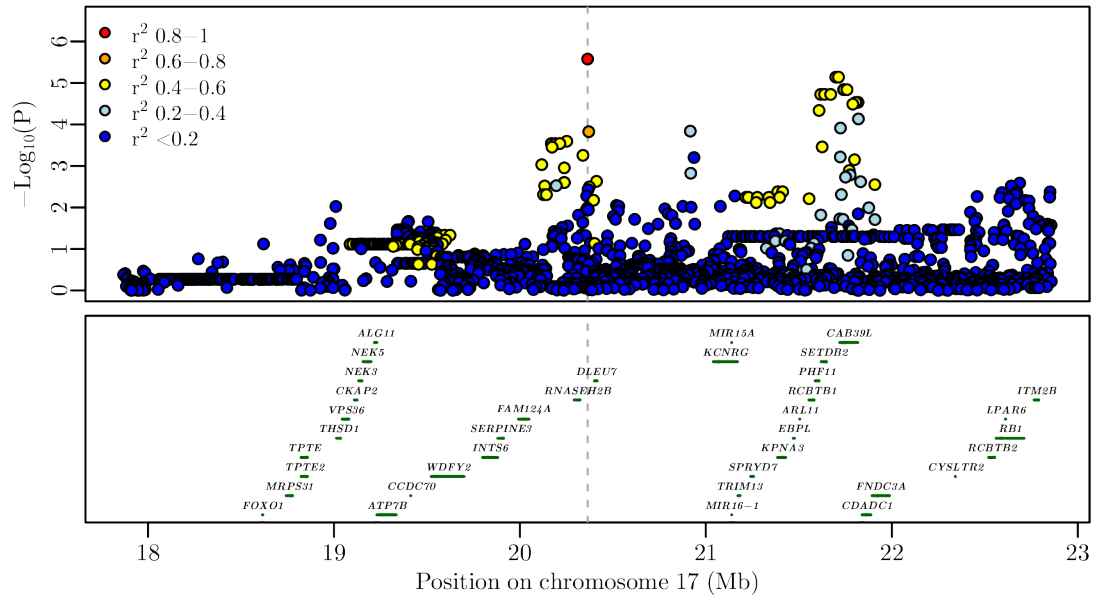
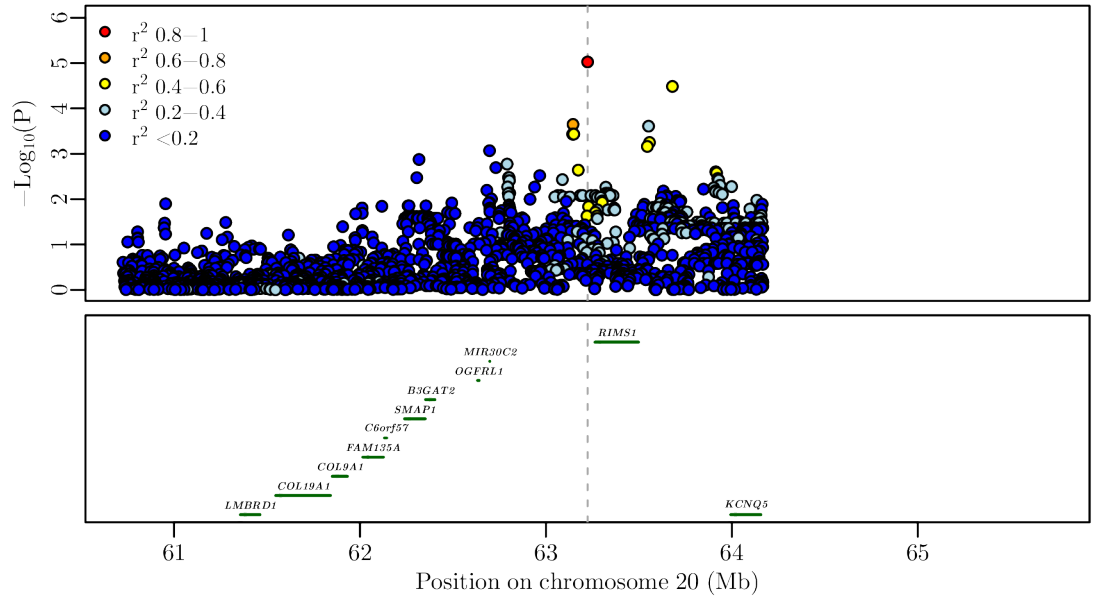


Figure D.5 *Continued from previous page*: Glucose 75-minute post-oral sugar test loci

(e)



(f)

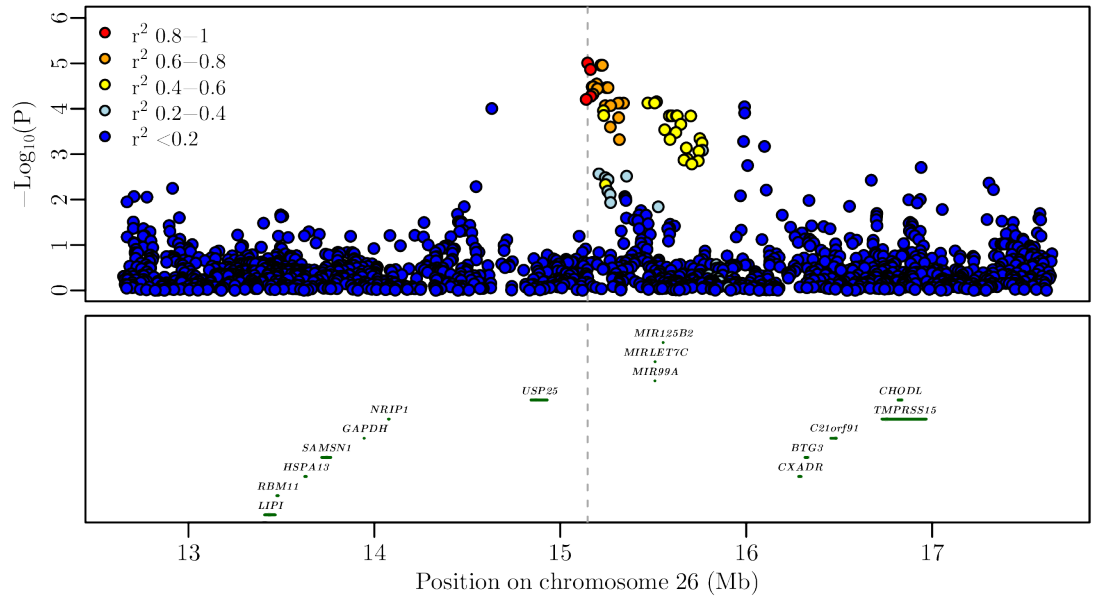
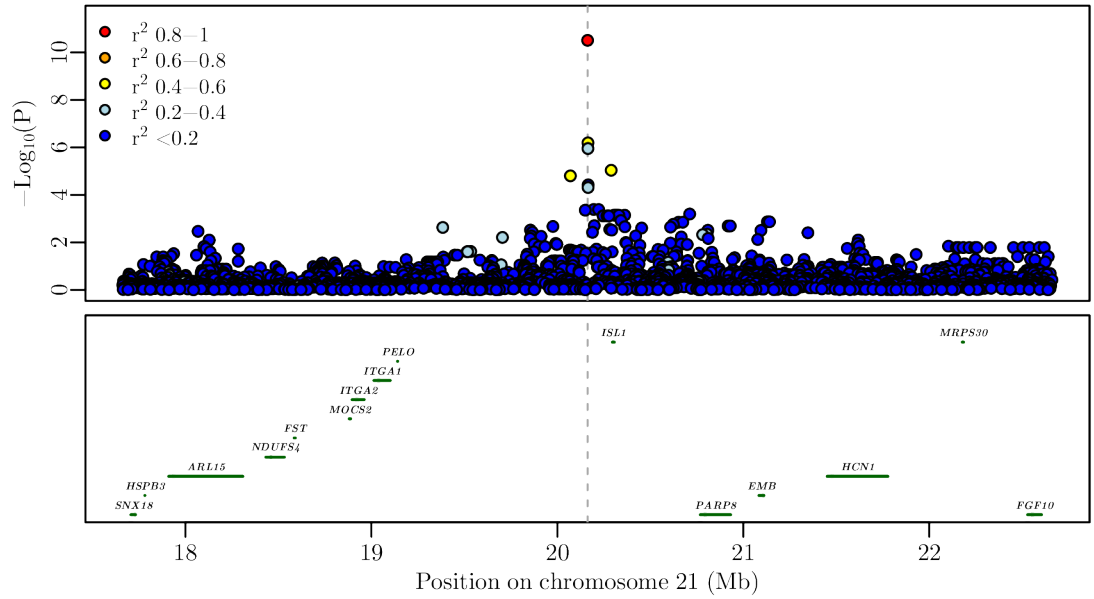


Figure D.6: Insulin 75-minute post-oral sugar test loci

(a)



(b)

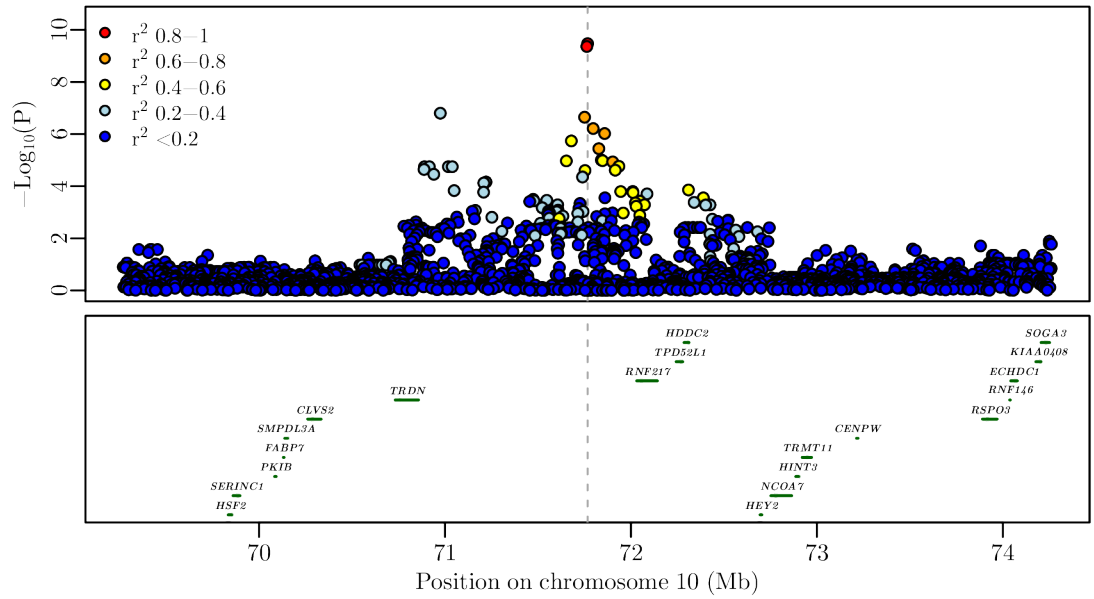
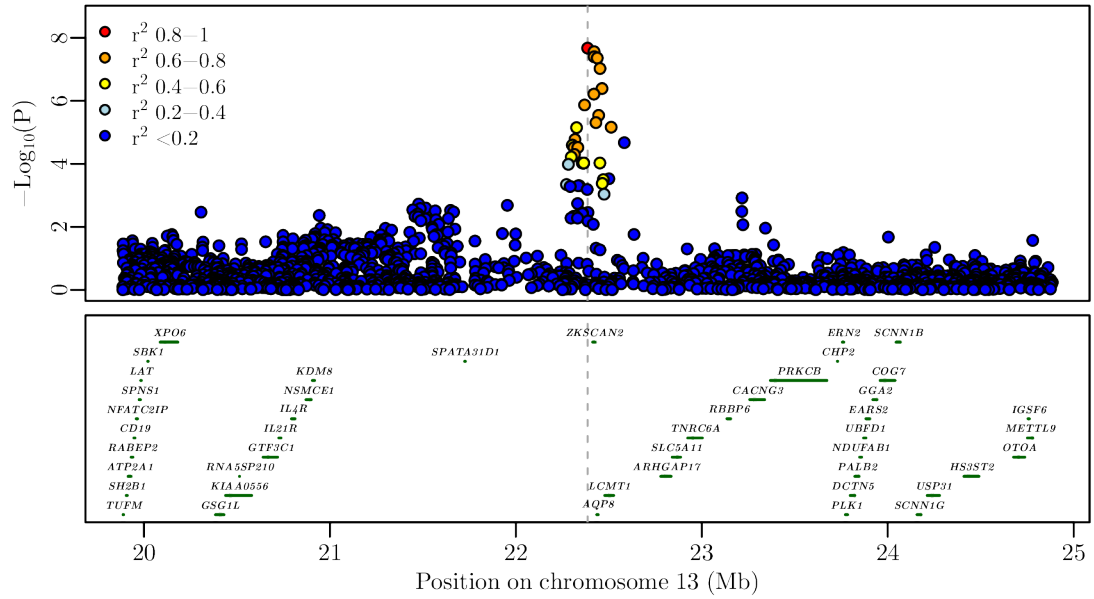


Figure D.6 *Continued from previous page*: Insulin 75-minute post-oral sugar test loci

(c)



(d)

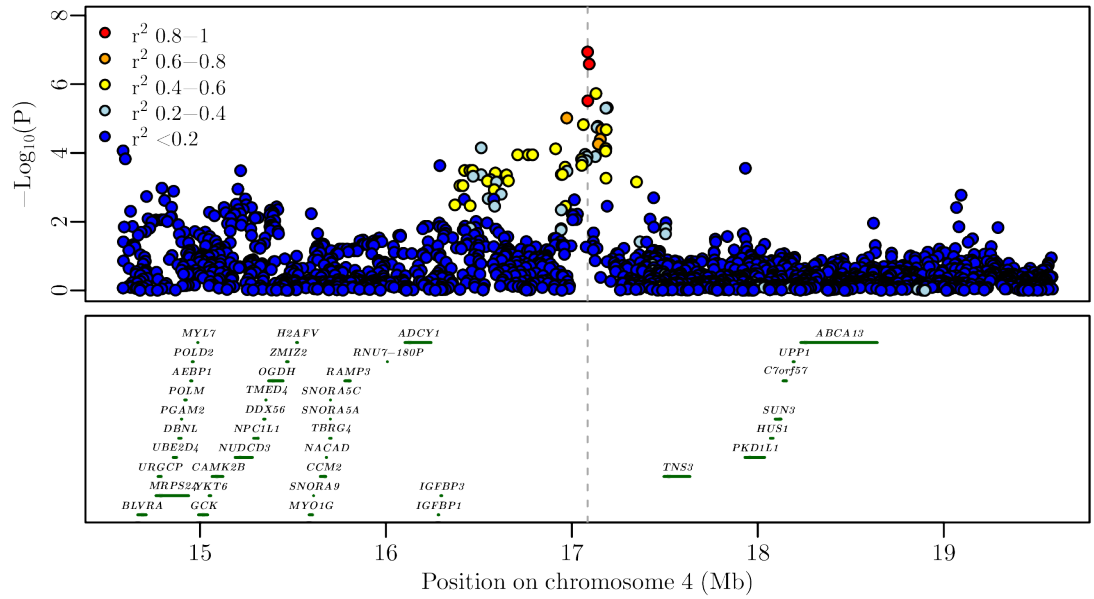
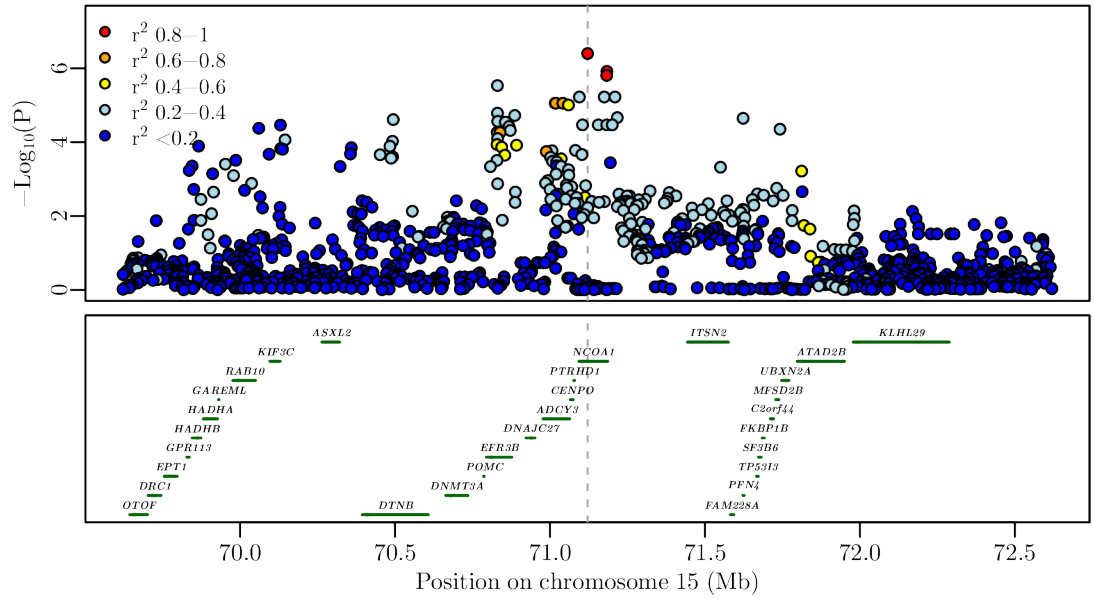


Figure D.6 *Continued from previous page*: Insulin 75-minute post-oral sugar test loci

(e)



(f)

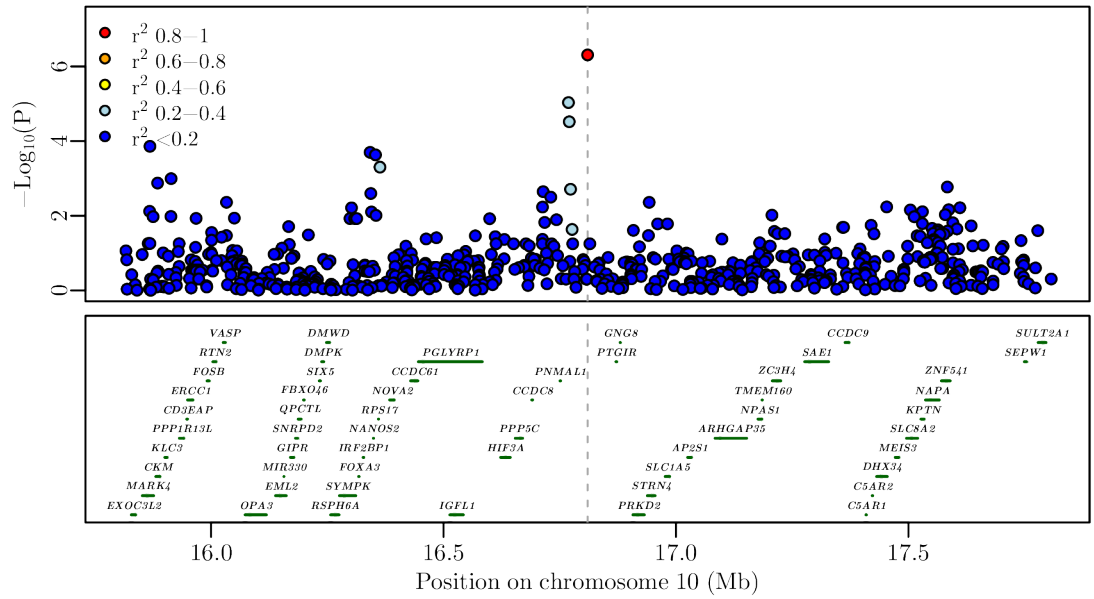
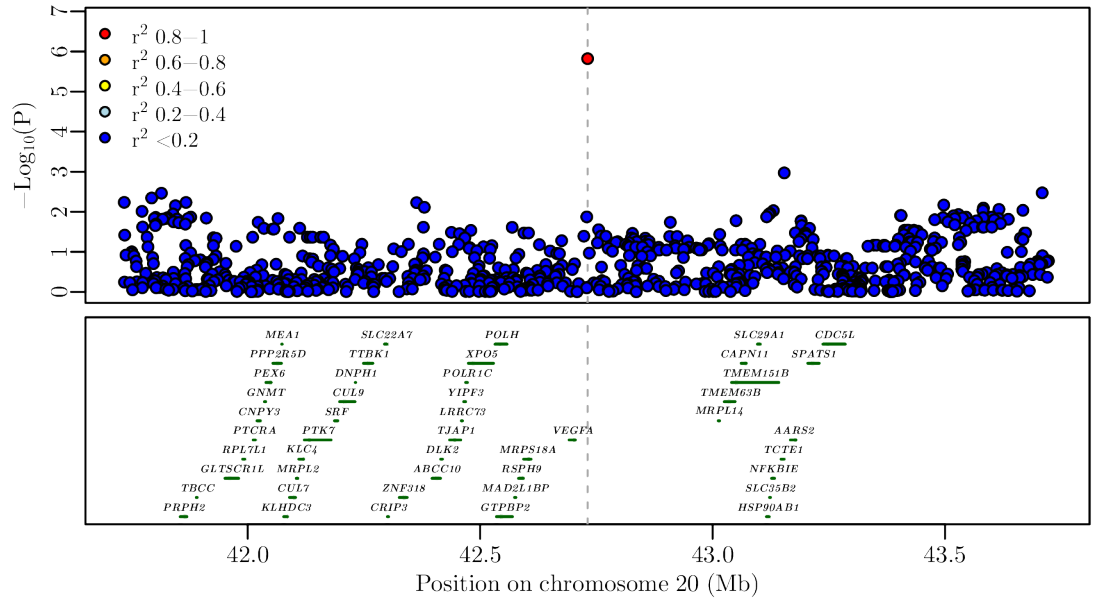


Figure D.6 *Continued from previous page*: Insulin 75-minute post-oral sugar test loci

(g)



(h)

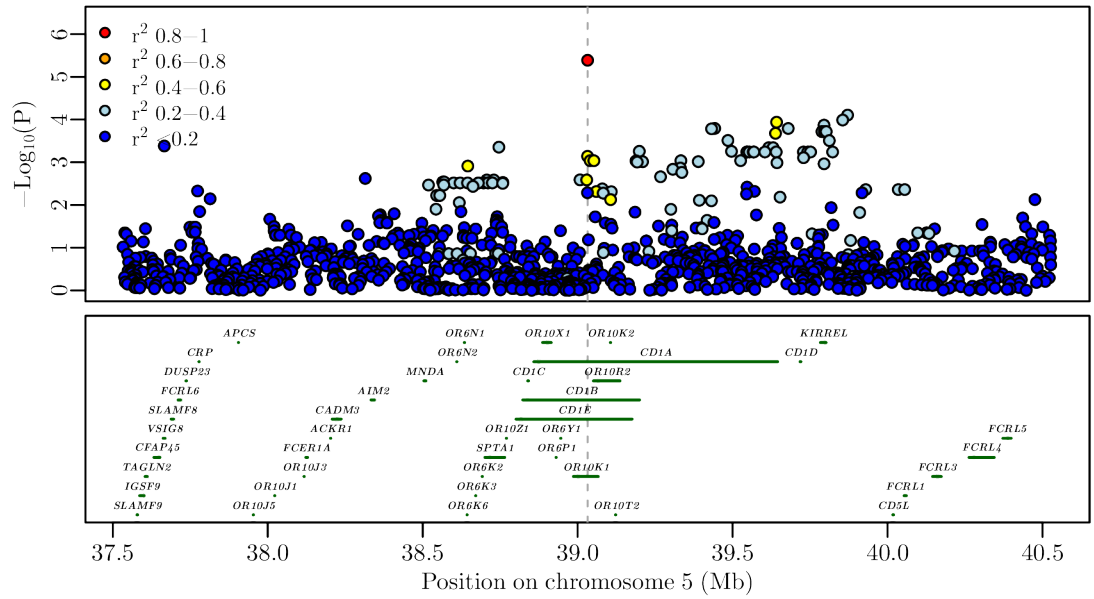
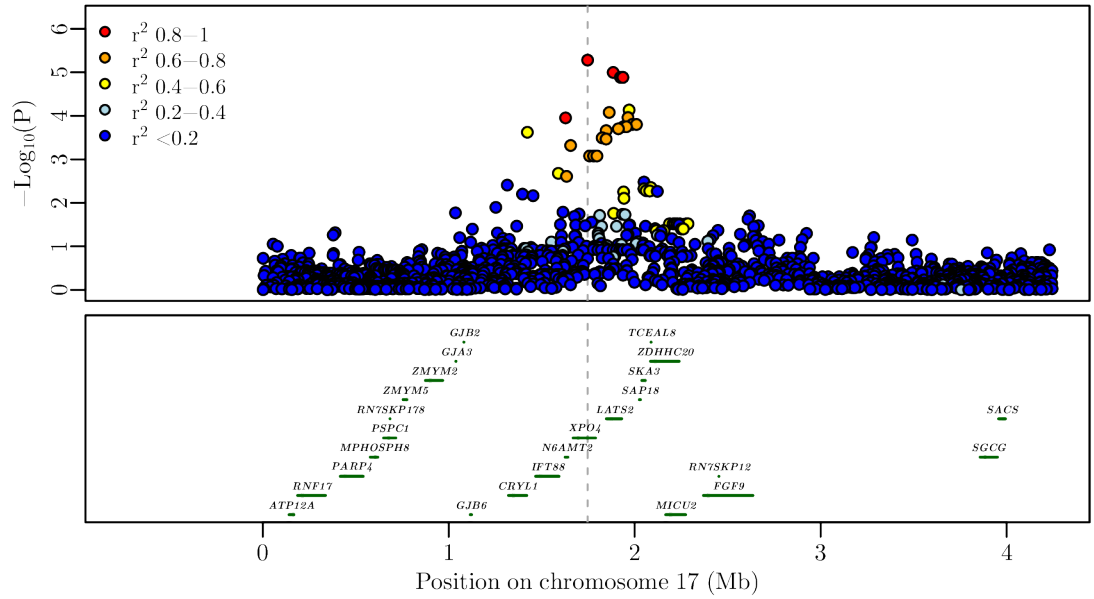


Figure D.6 *Continued from previous page*: Insulin 75-minute post-oral sugar test loci

(i)



(j)

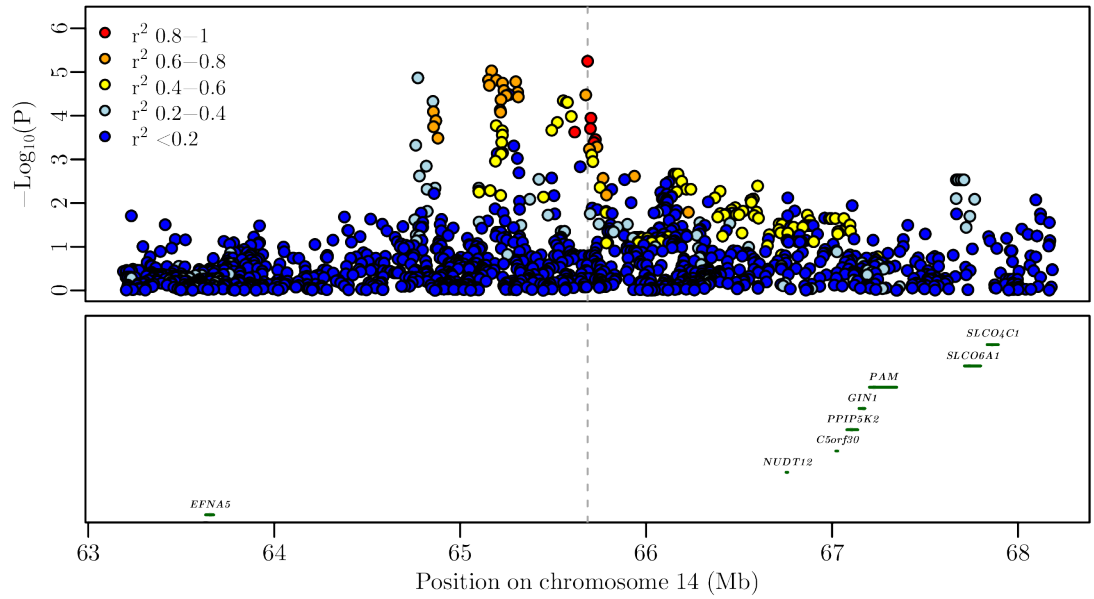


Figure D.6 *Continued from previous page*: Insulin 75-minute post-oral sugar test loci

(k)

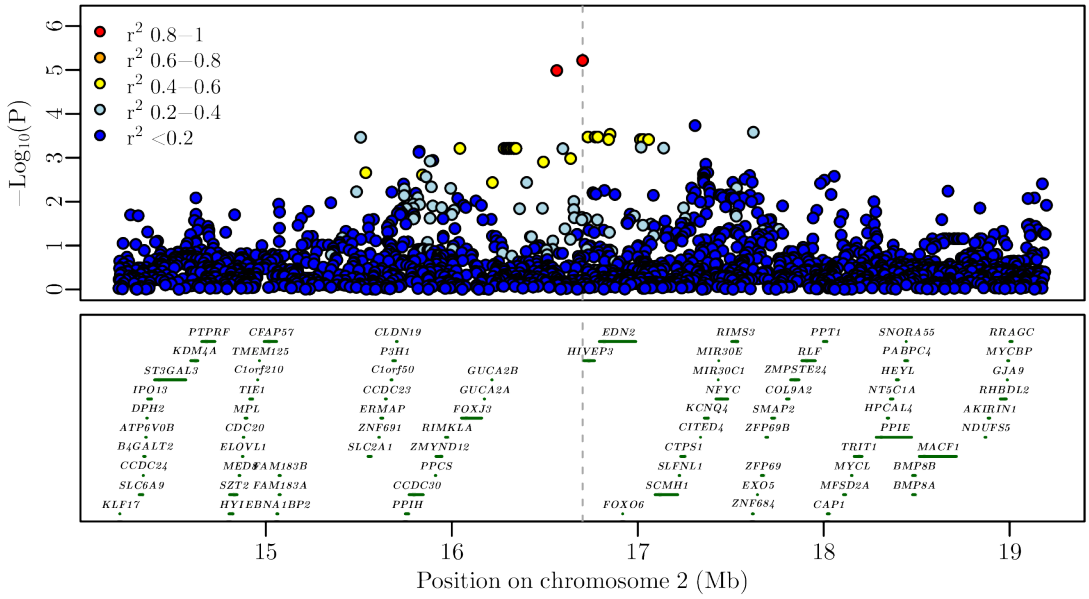
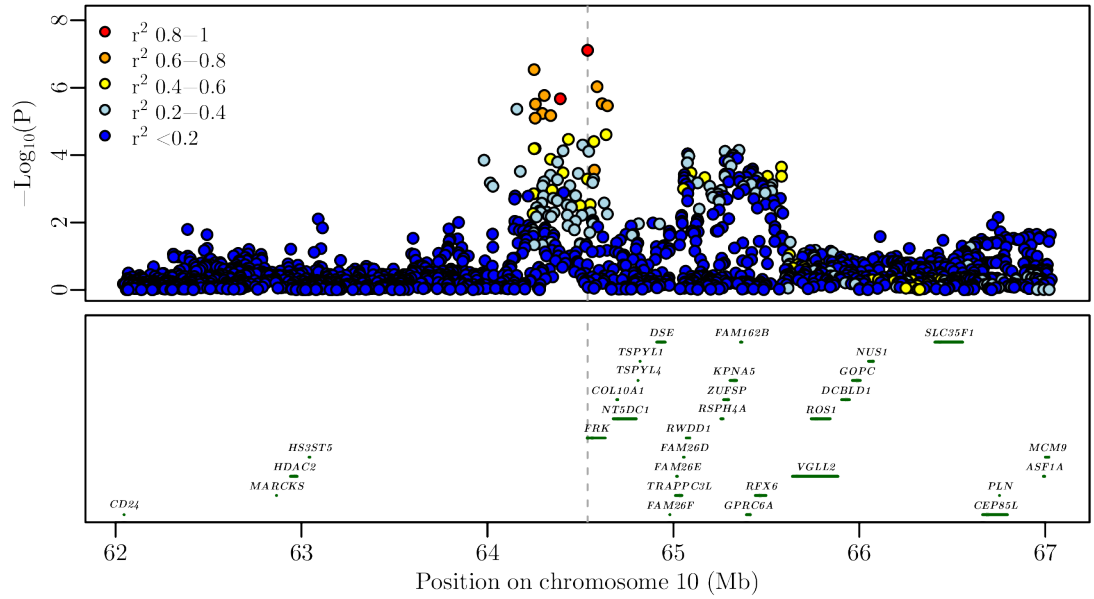


Figure D.7: Triglyceride loci

(a)



(b)

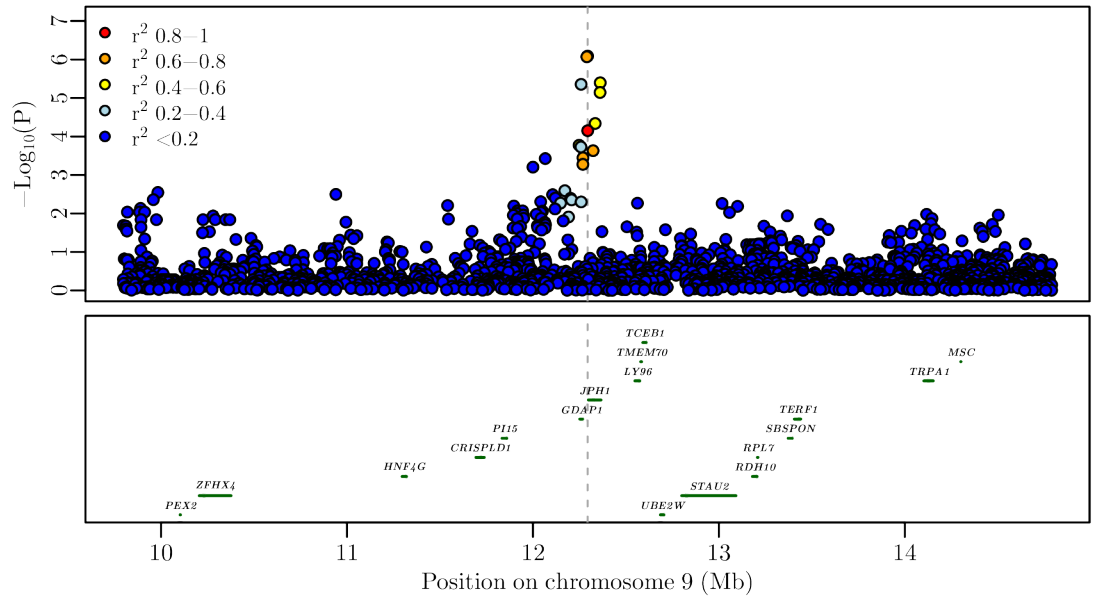
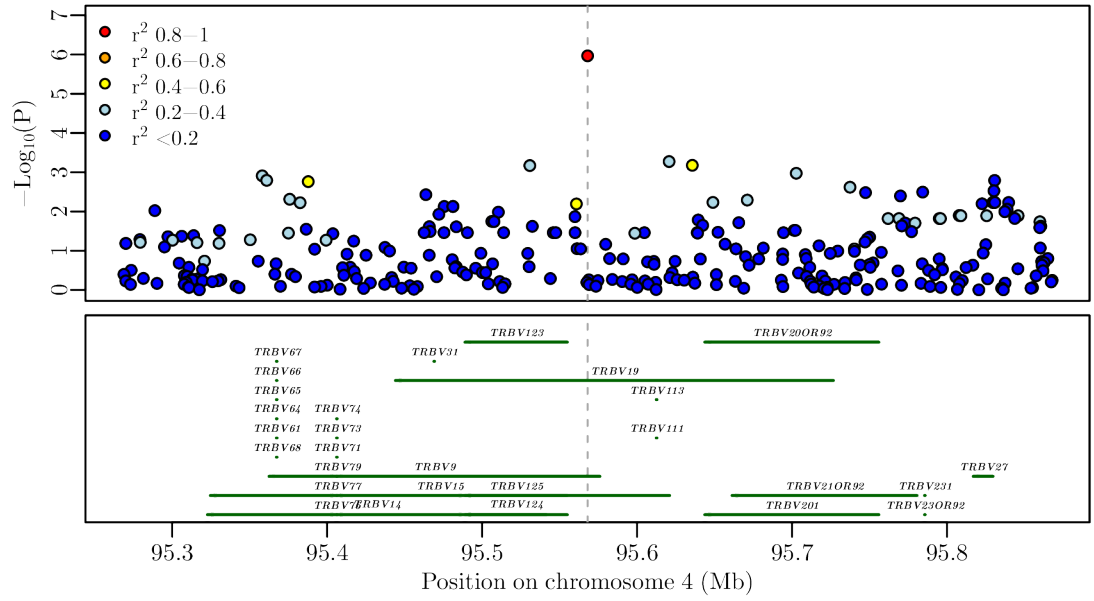


Figure D.7 *Continued from previous page*: Triglyceride loci

(c)



(d)

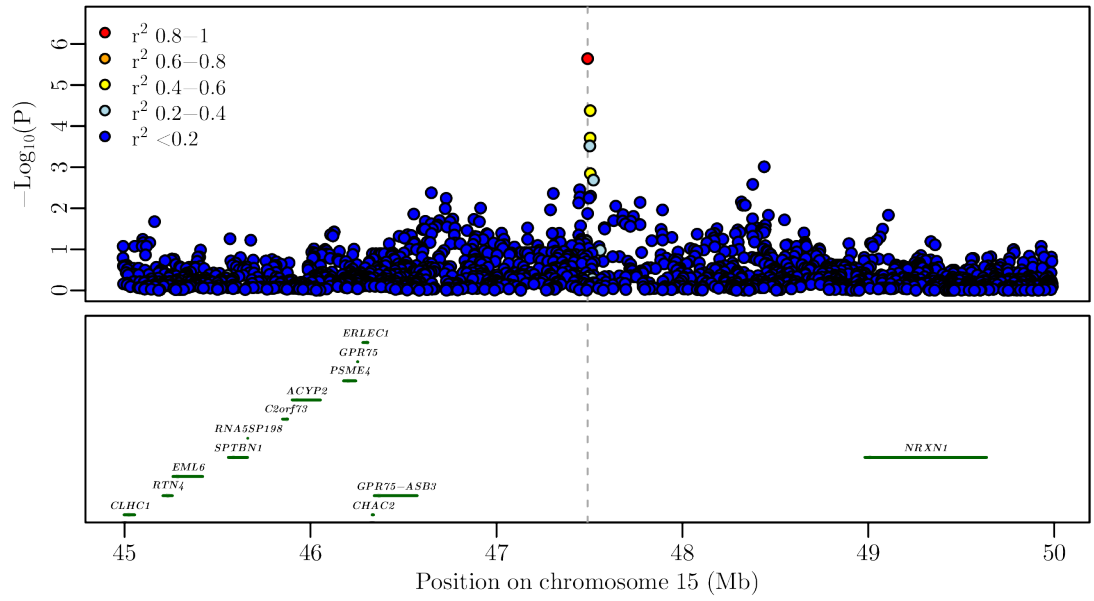
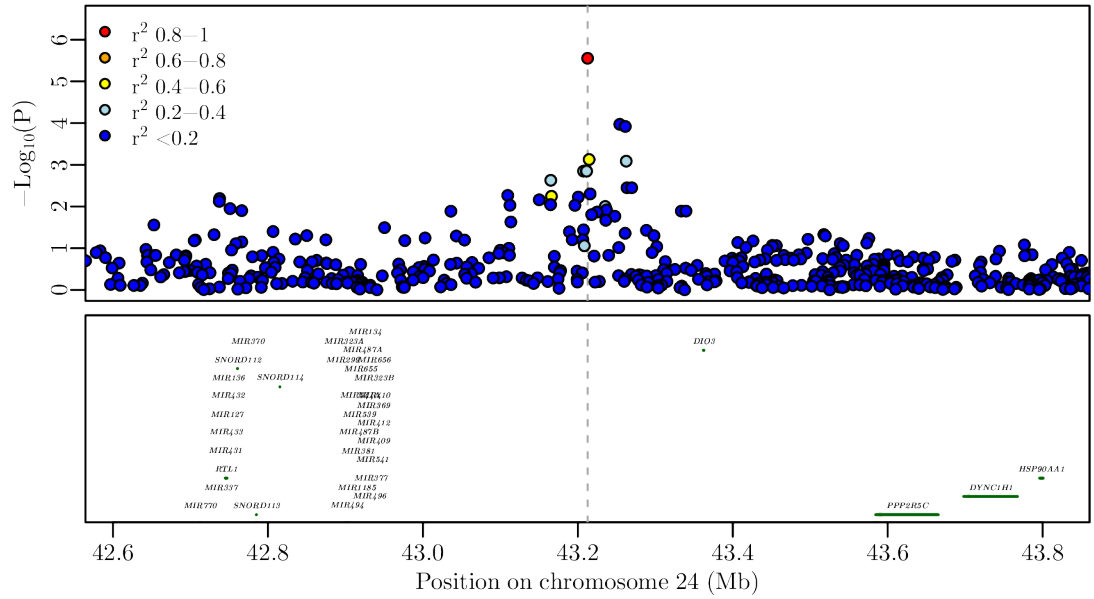


Figure D.7 *Continued from previous page*: Triglyceride loci

(e)



(f)

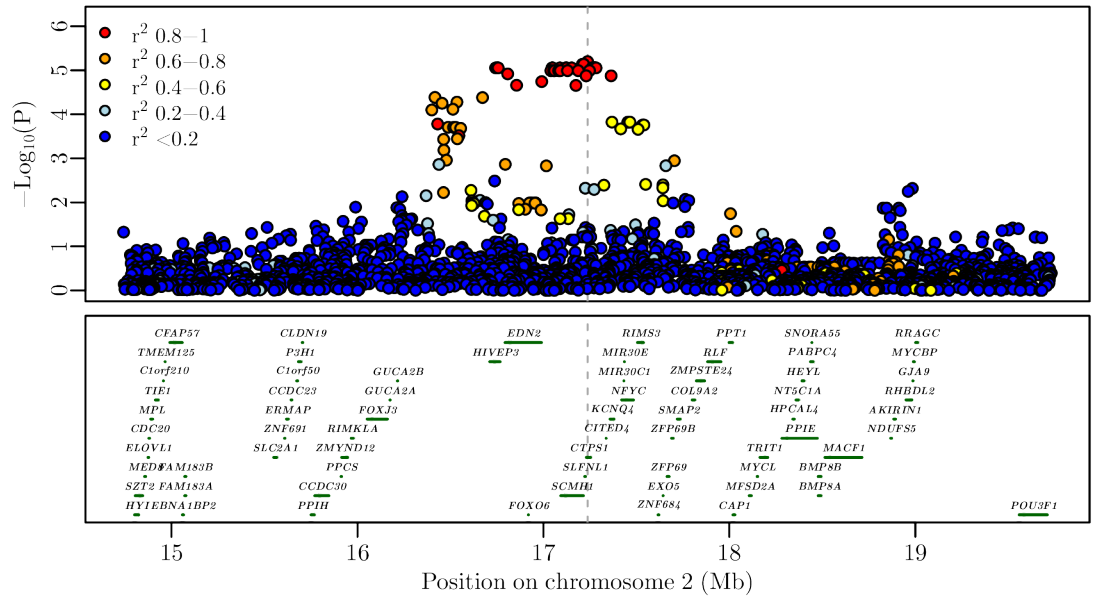


Figure D.7 *Continued from previous page*: Triglyceride loci

(e)

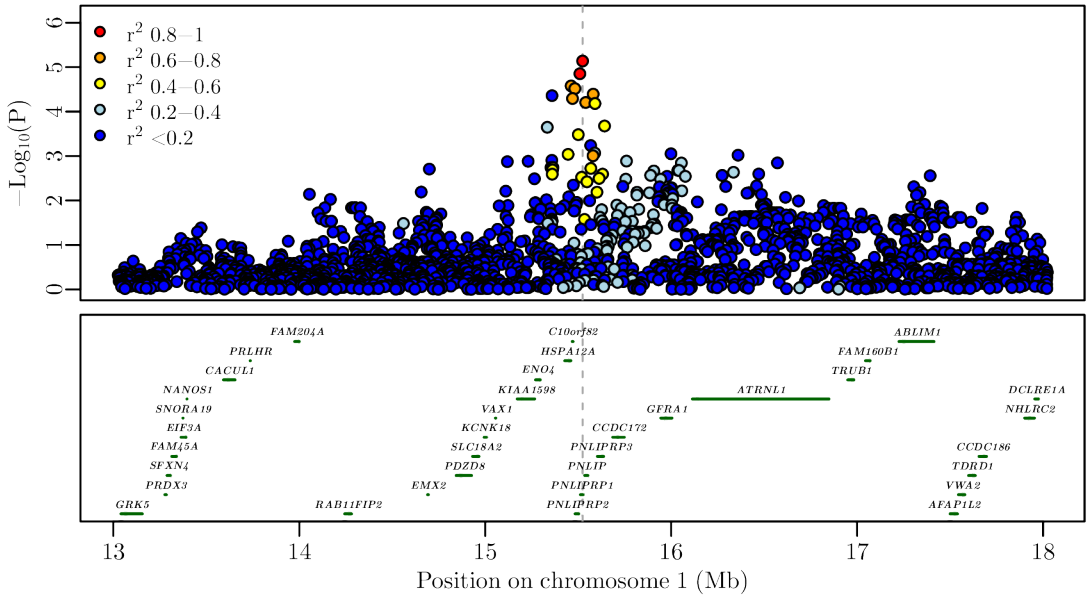
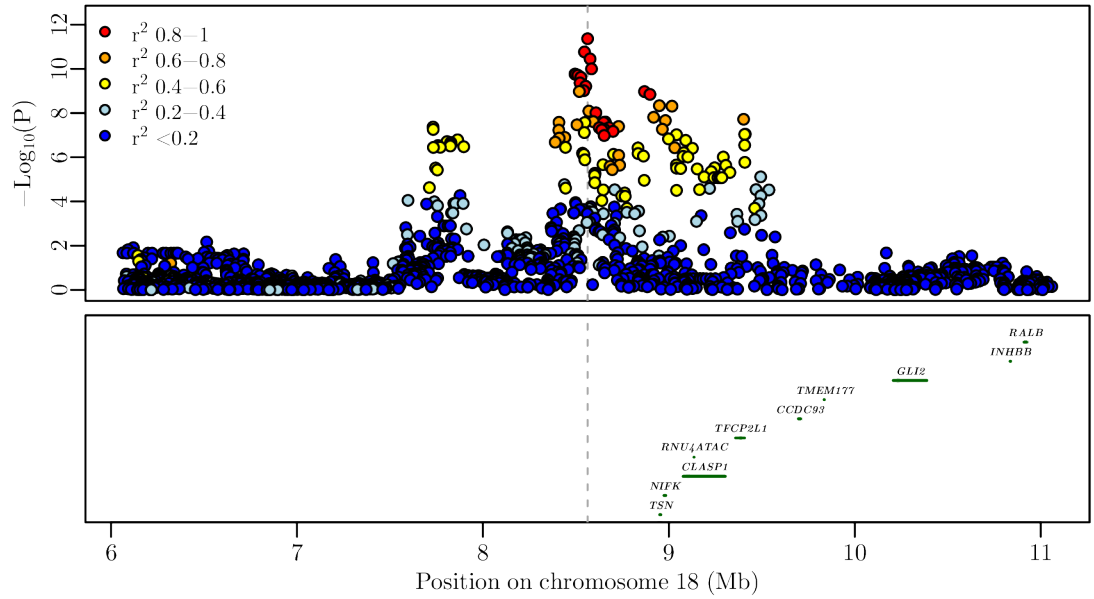


Figure D.8: Non-esterified fatty acid loci

(a)



(b)

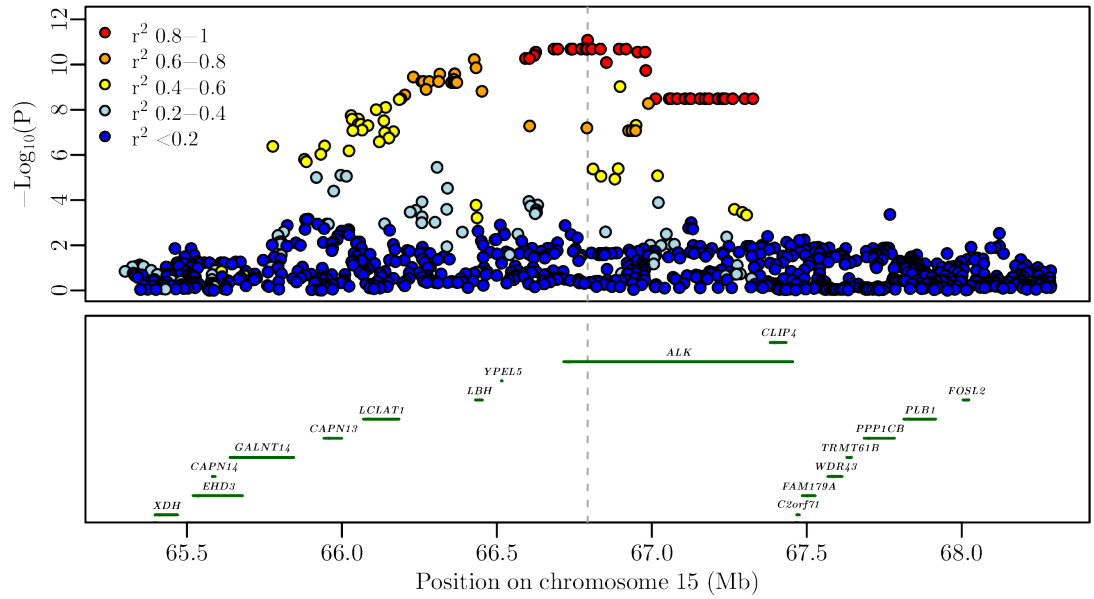
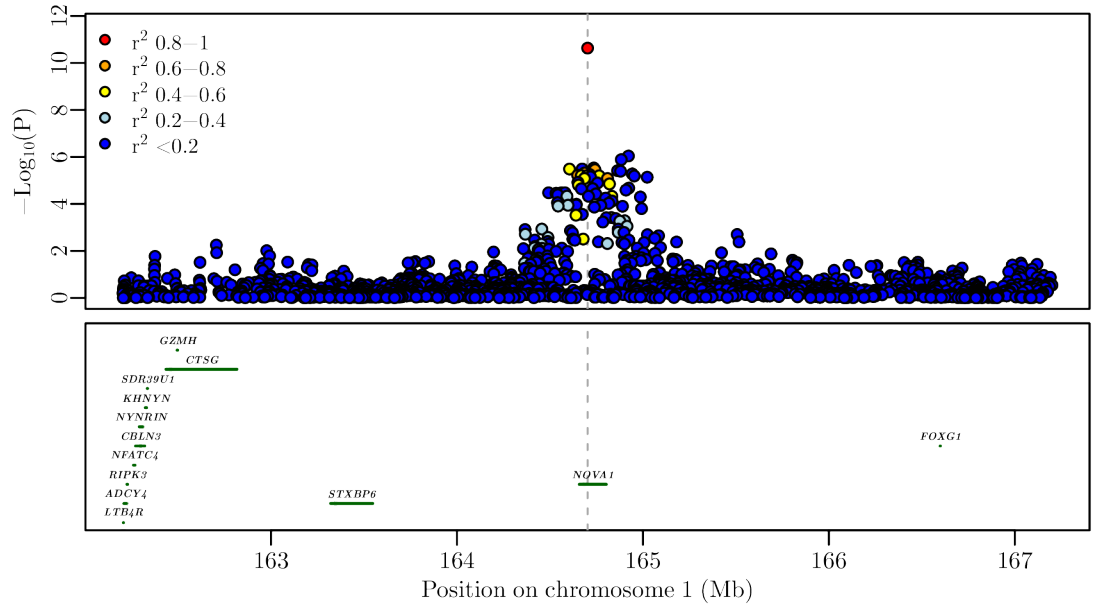


Figure D.8 *Continued from previous page*: Non-esterified fatty acid loci

(c)



(d)

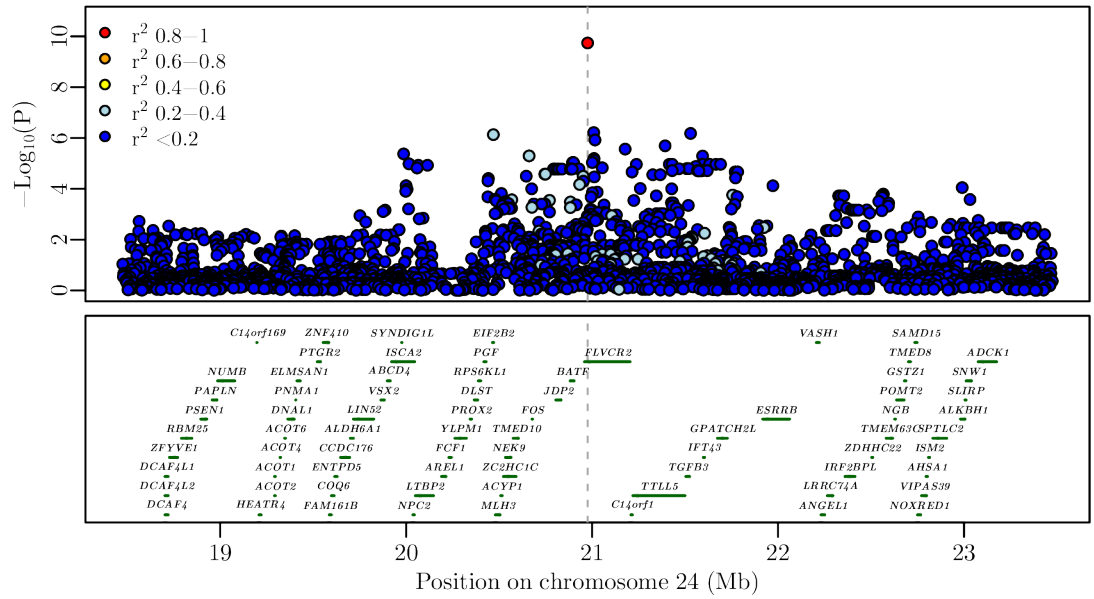
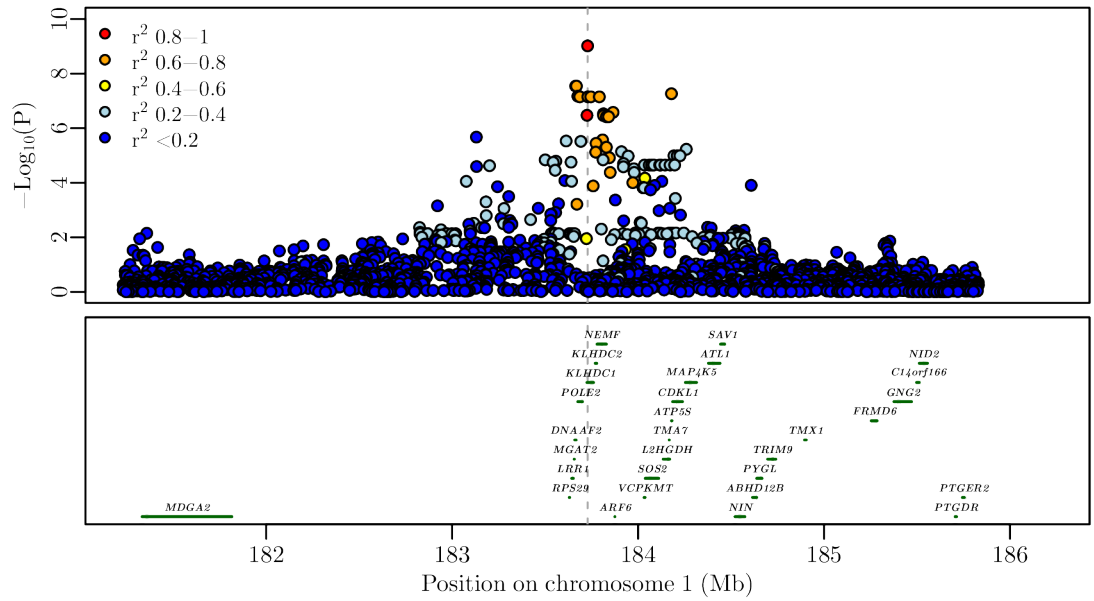


Figure D.8 *Continued from previous page*: Non-esterified fatty acid loci

(e)



(f)

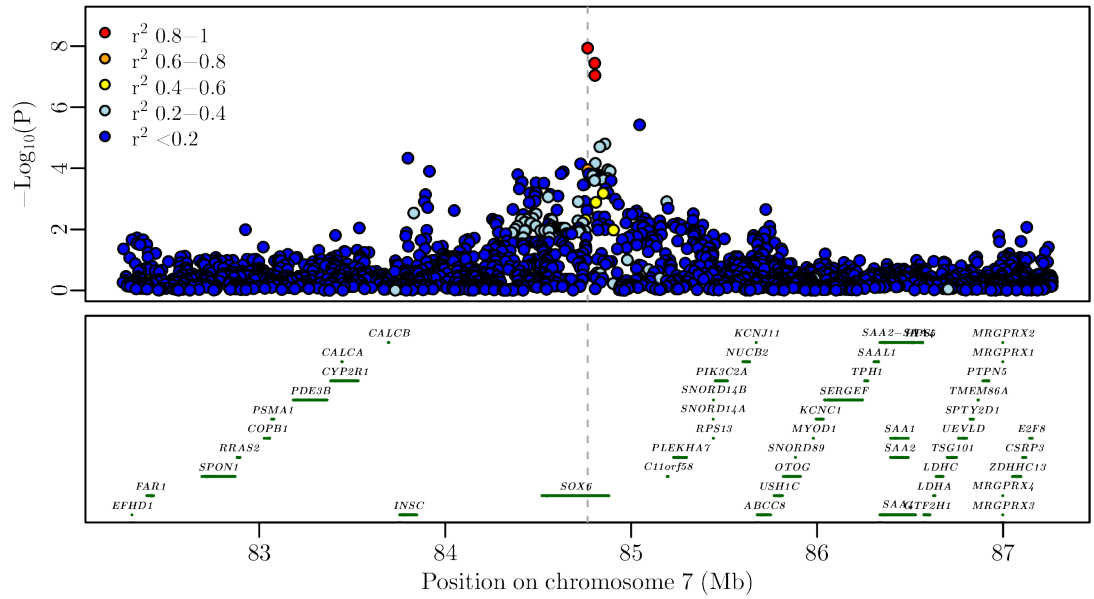
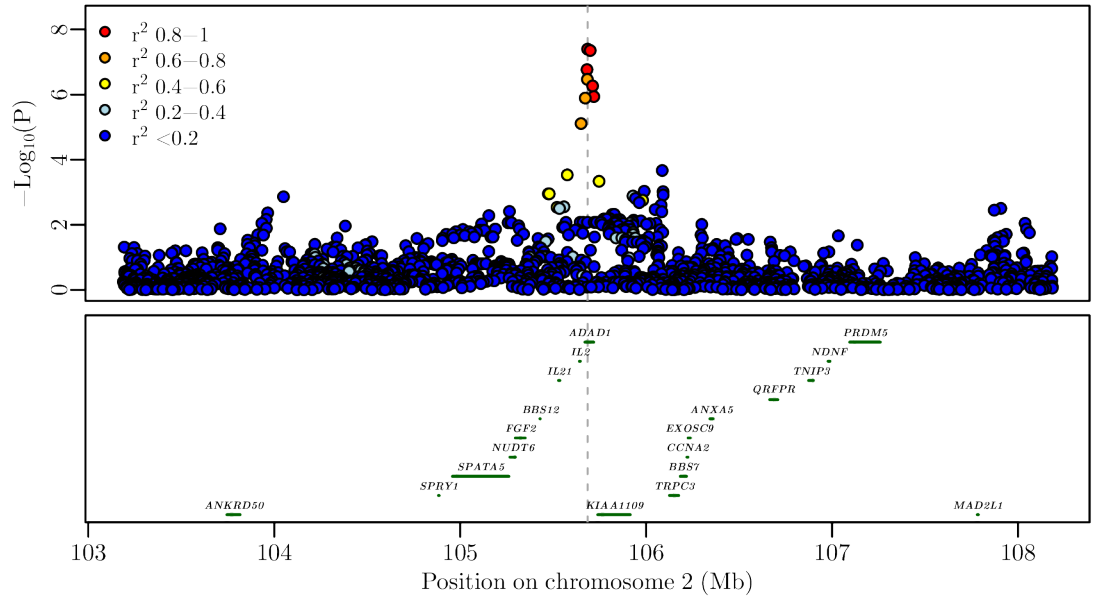


Figure D.8 *Continued from previous page*: Non-esterified fatty acid loci

(g)



(h)

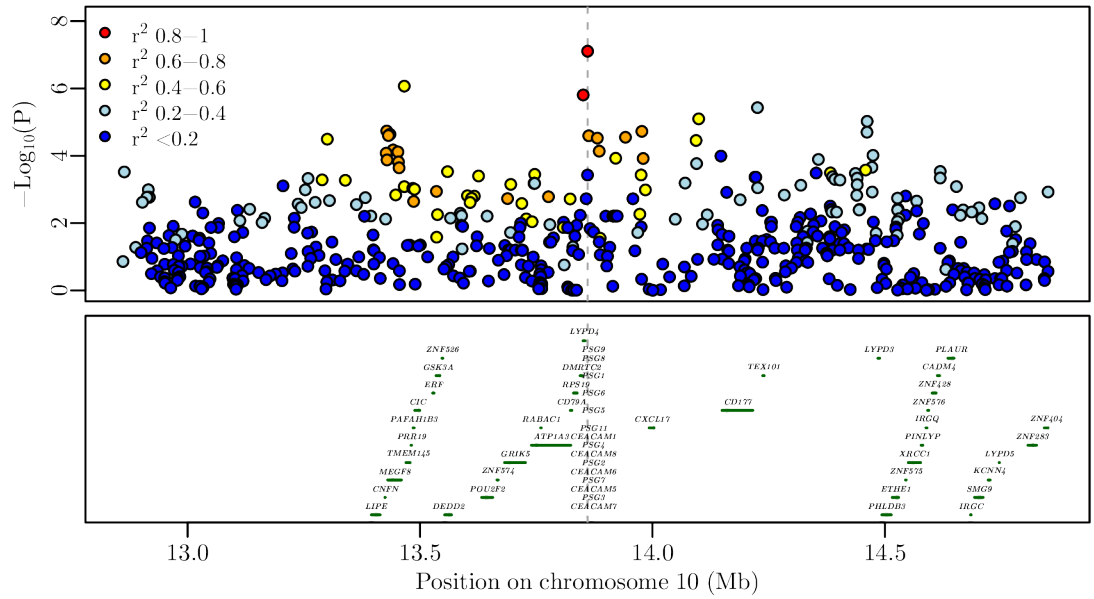
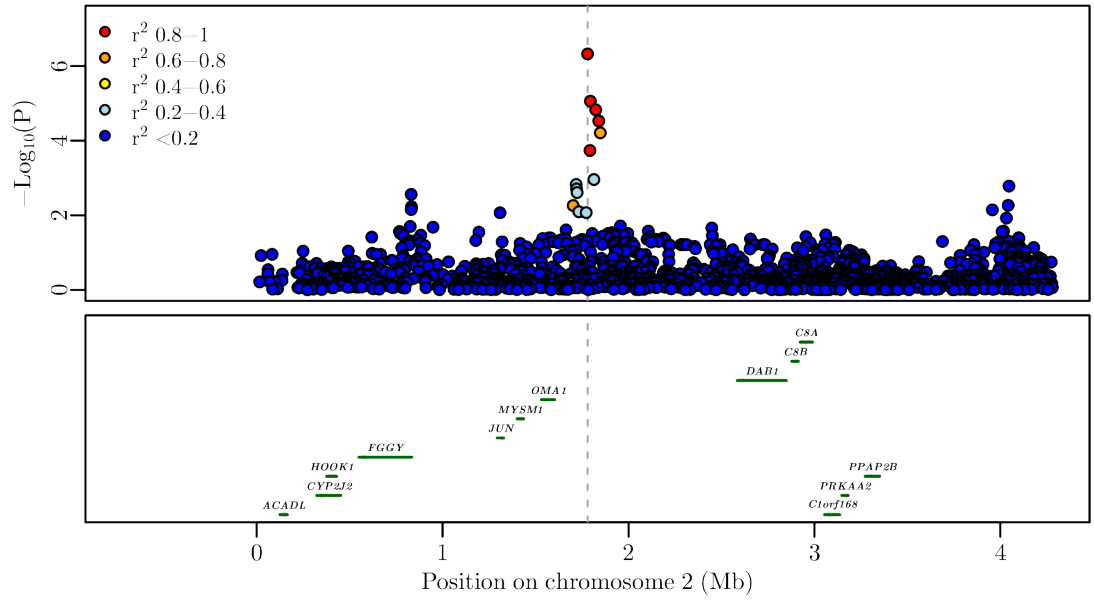


Figure D.8 *Continued from previous page*: Non-esterified fatty acid loci

(i)



(j)

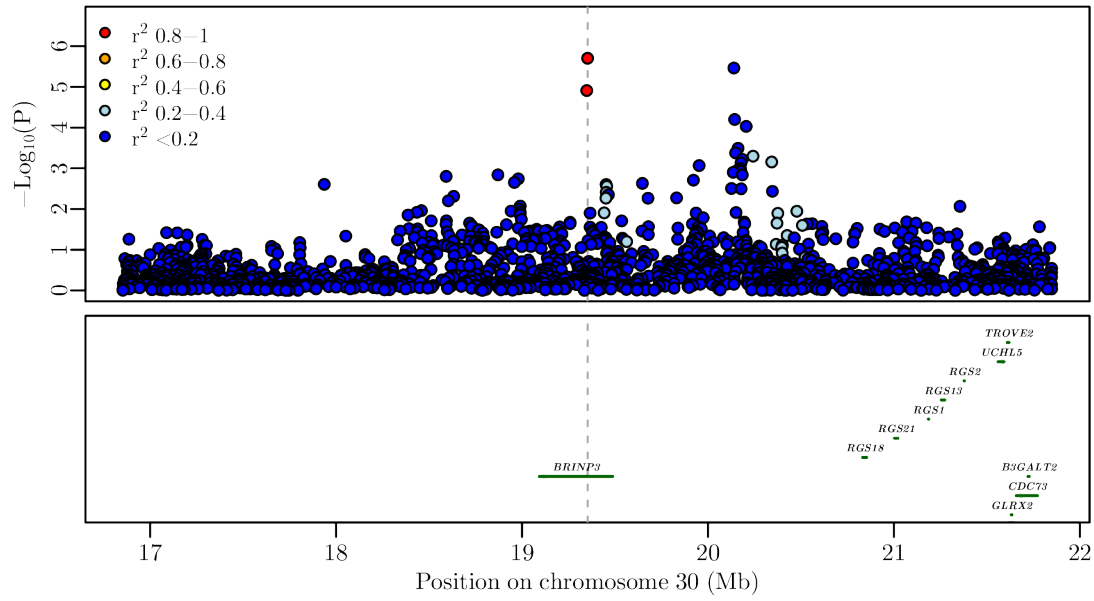
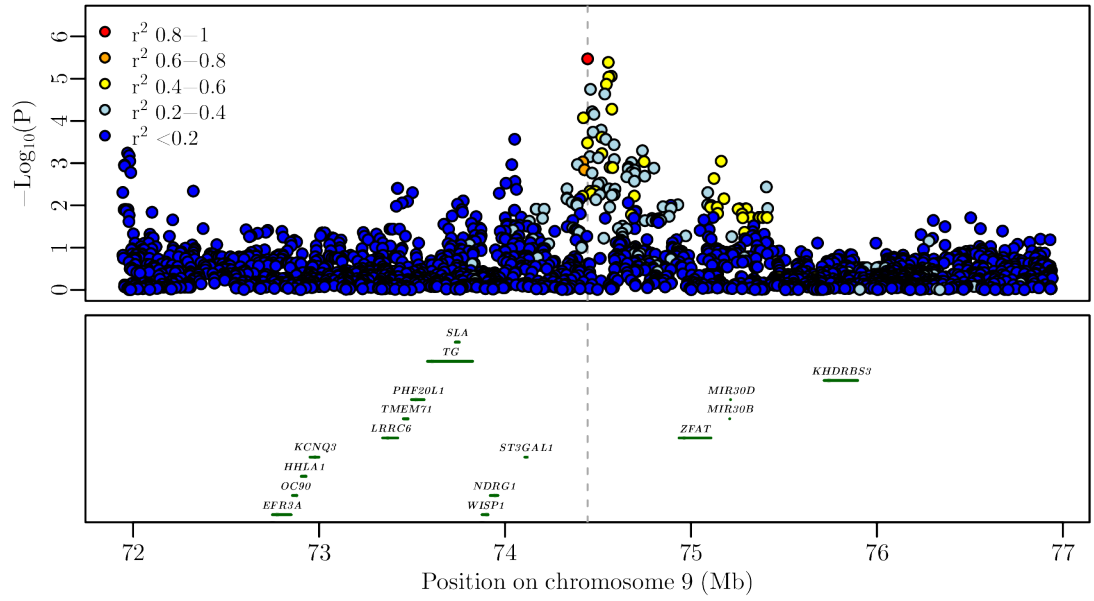


Figure D.8 *Continued from previous page*: Non-esterified fatty acid loci

(k)



(l)

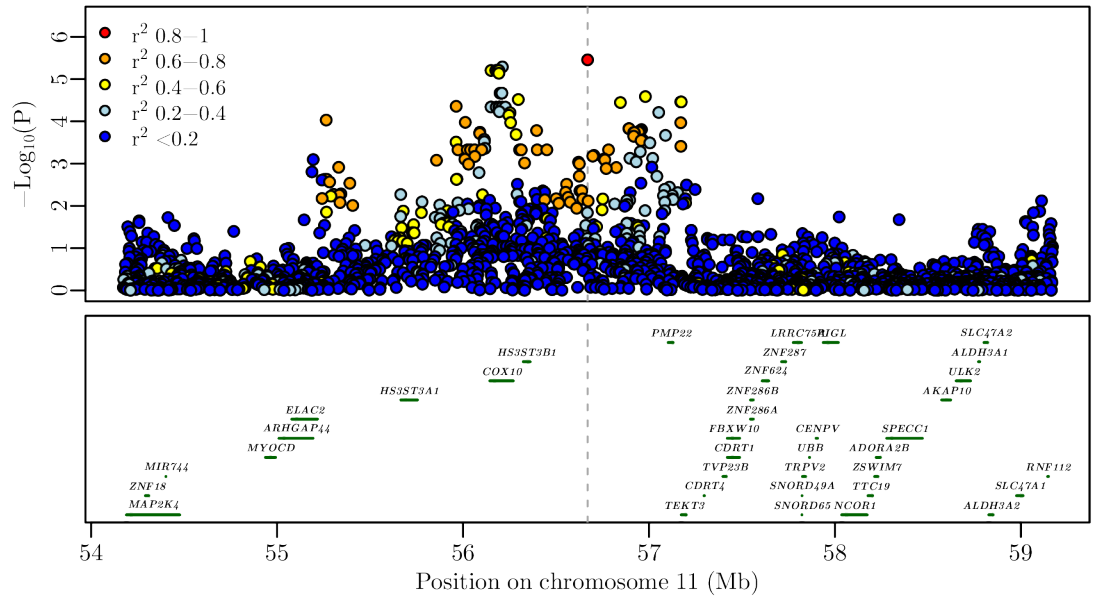
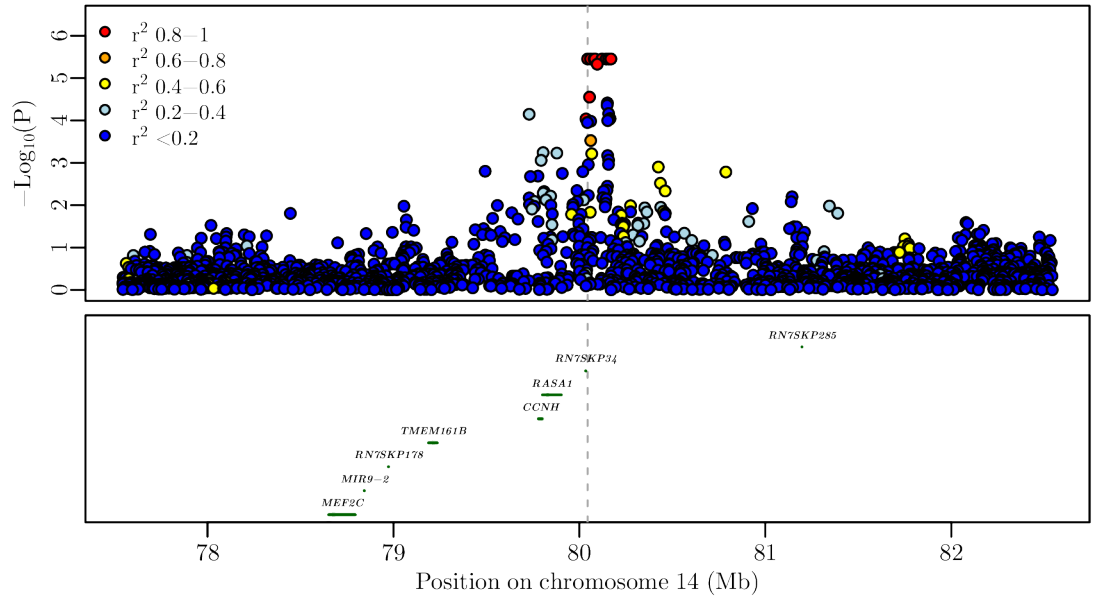


Figure D.8 *Continued from previous page*: Non-esterified fatty acid loci

(m)



(n)

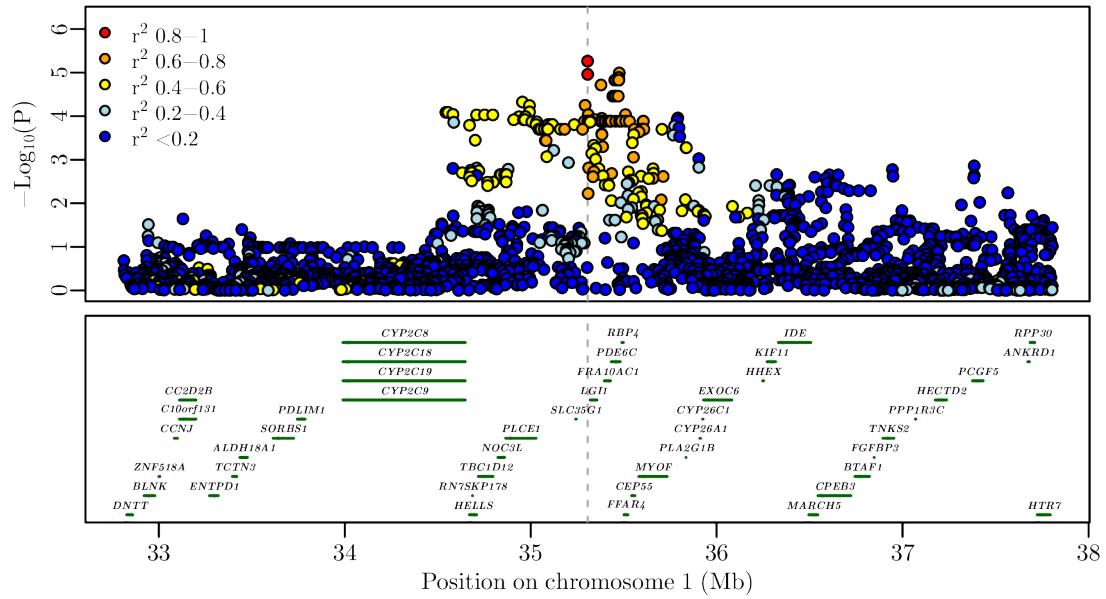
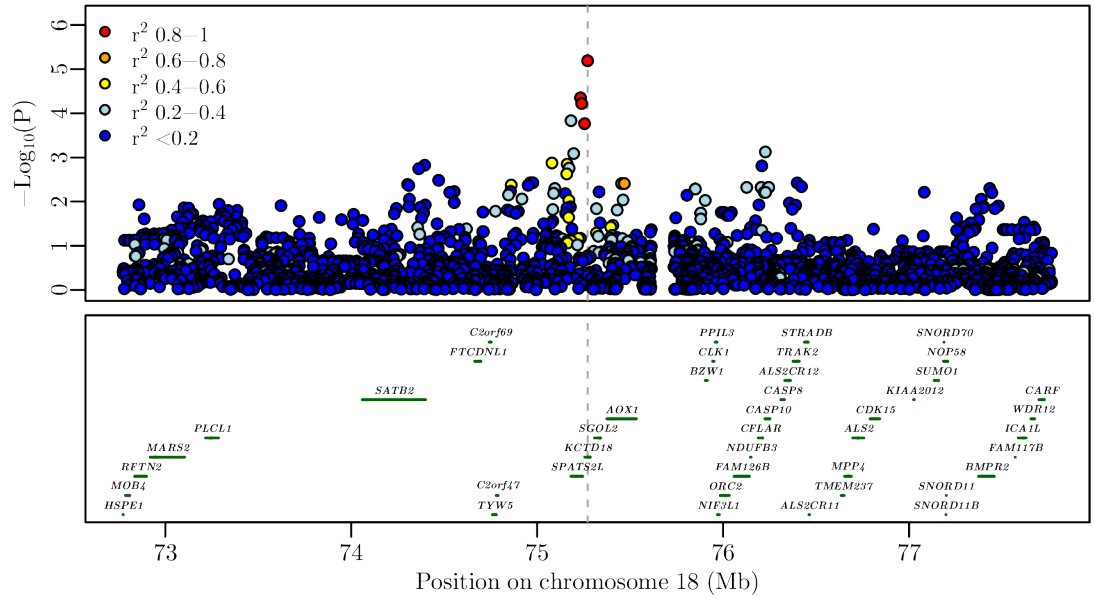


Figure D.8 *Continued from previous page*: Non-esterified fatty acid loci

(o)



(p)

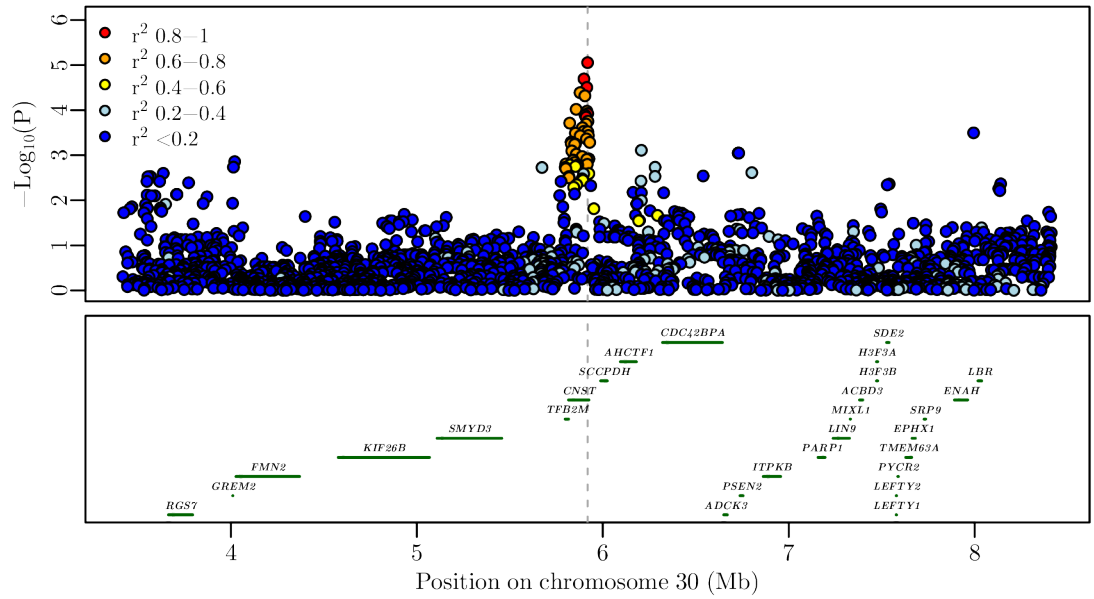
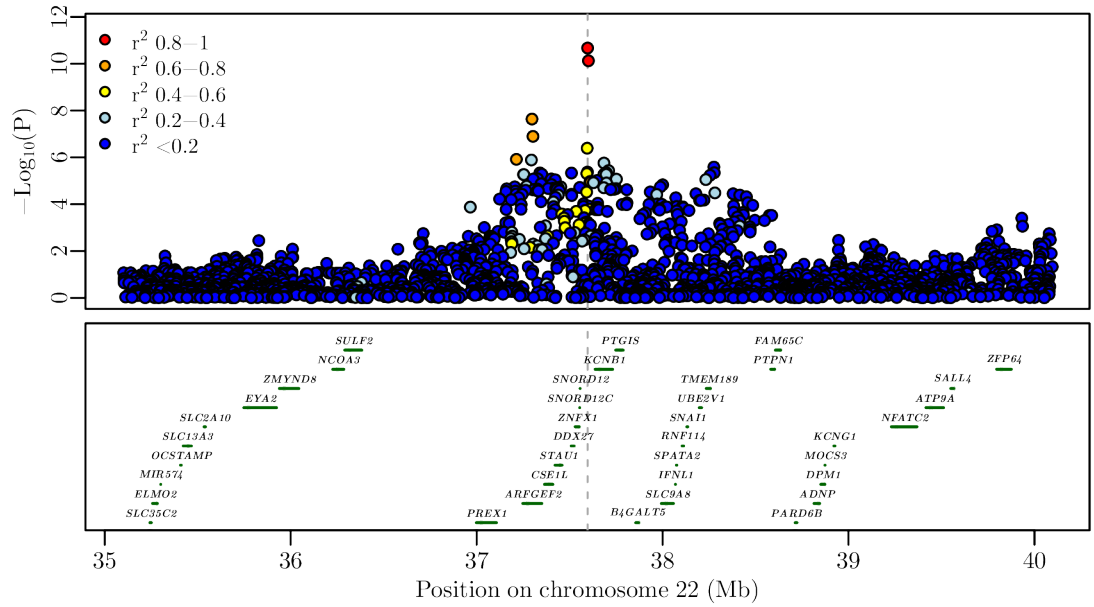


Figure D.9: Adrenocorticotropin hormone loci

(a)



(b)

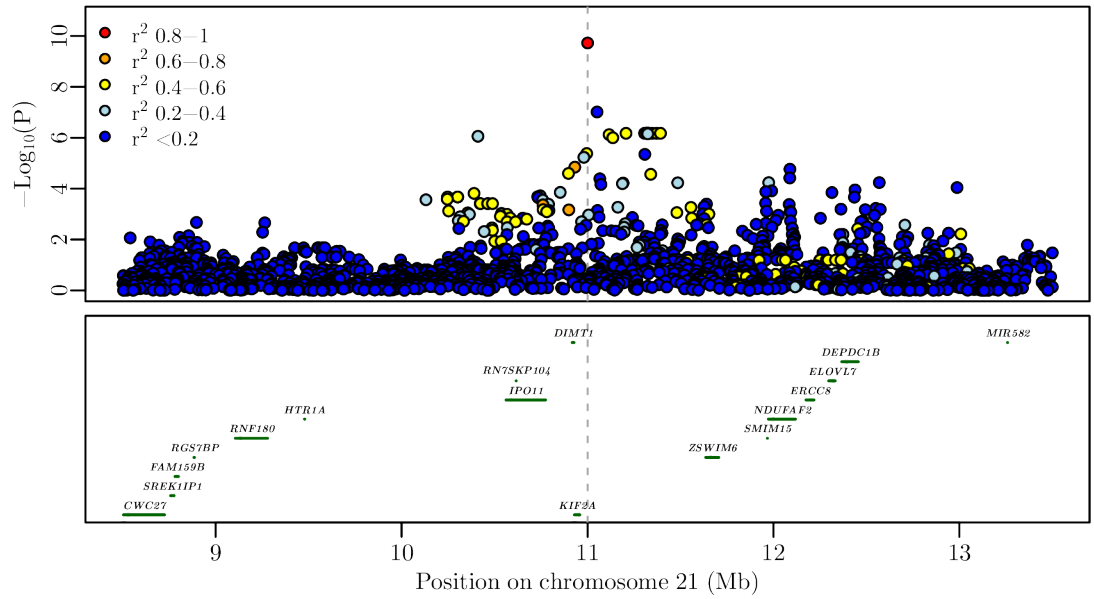
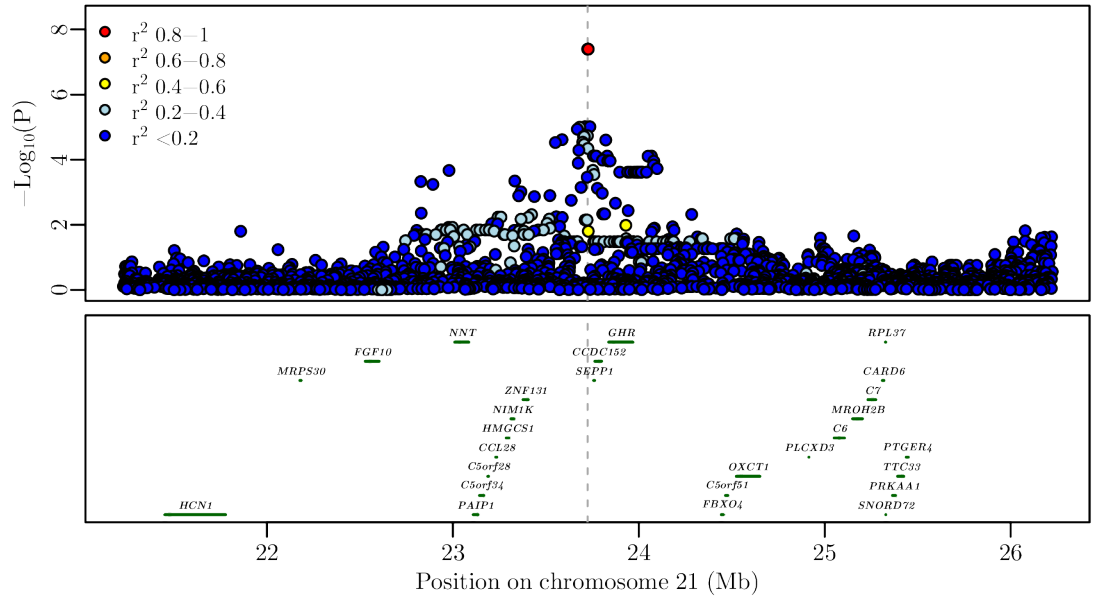


Figure D.9 *Continued from previous page*: Adrenocorticotropin hormone loci

(c)



(d)

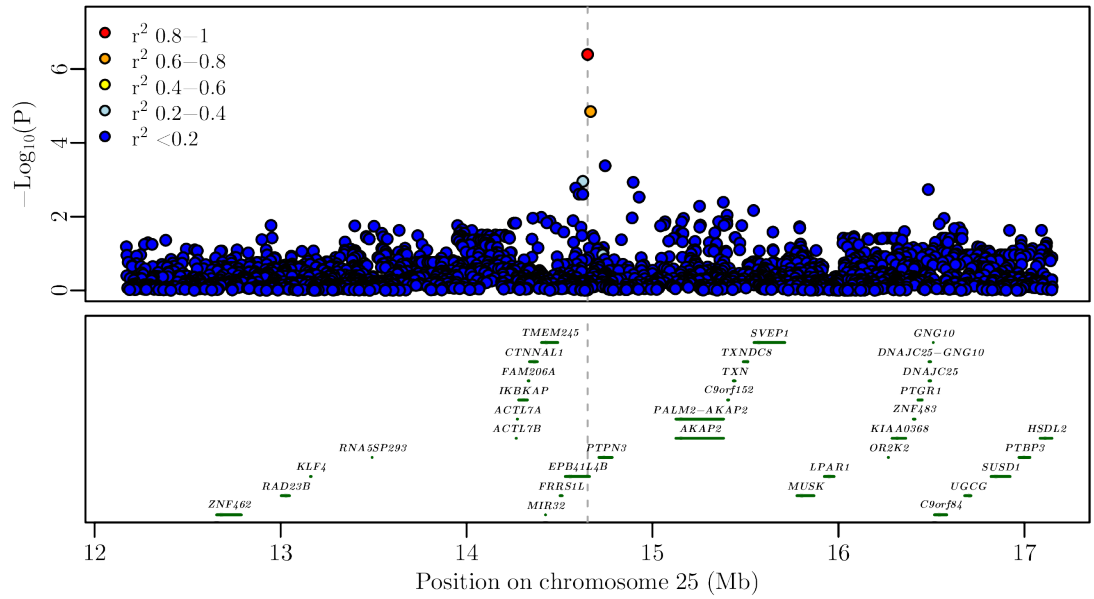
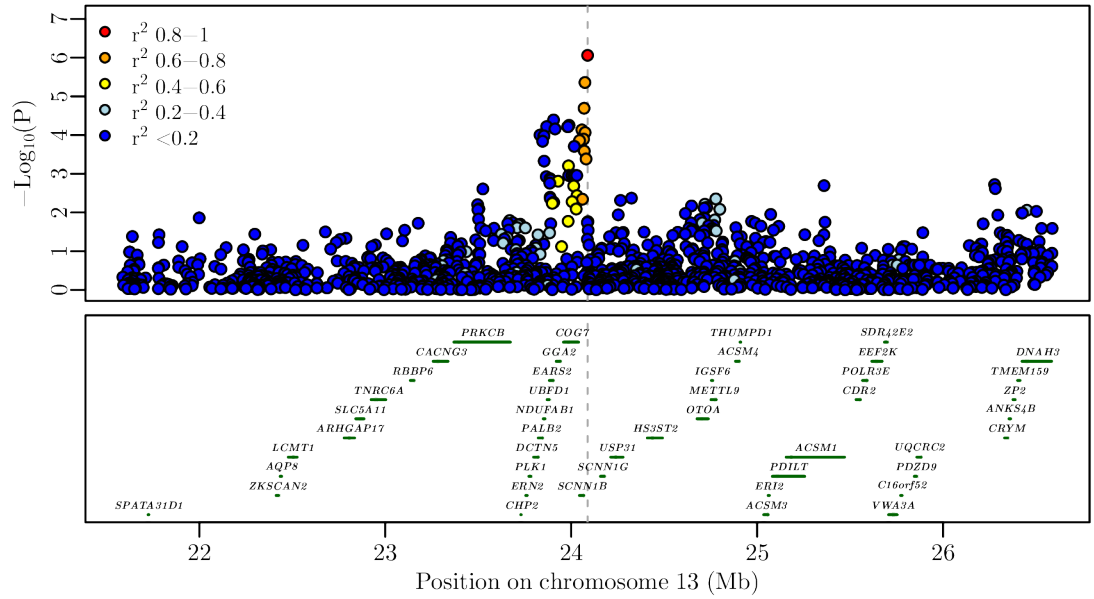


Figure D.9 *Continued from previous page*: Adrenocorticotropin hormone loci

(e)



(f)

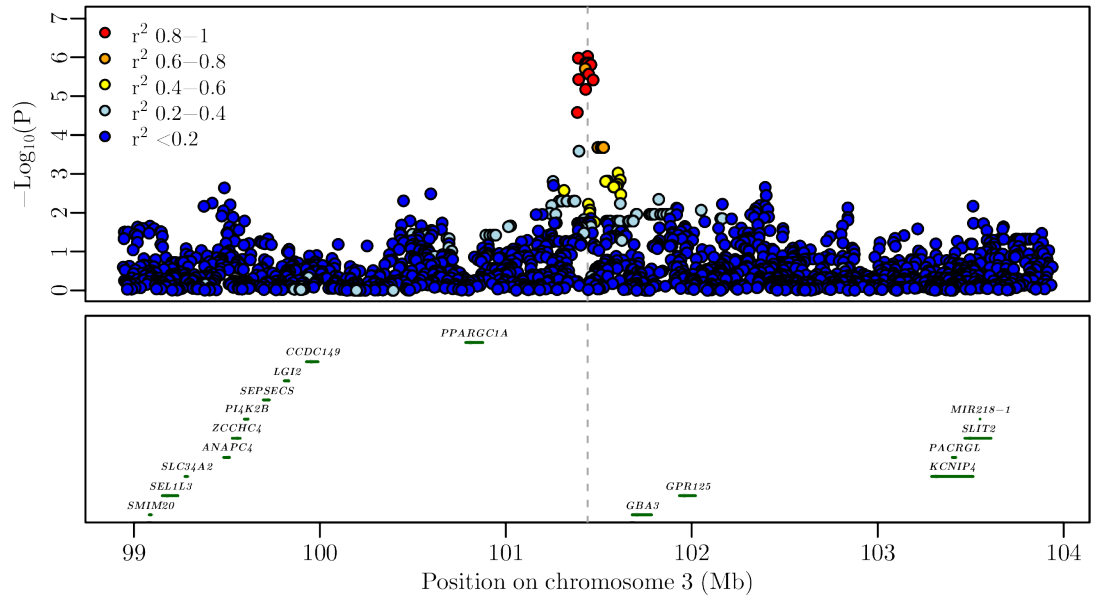


Figure D.9 *Continued from previous page*: Adrenocorticotropin hormone loci

(g)

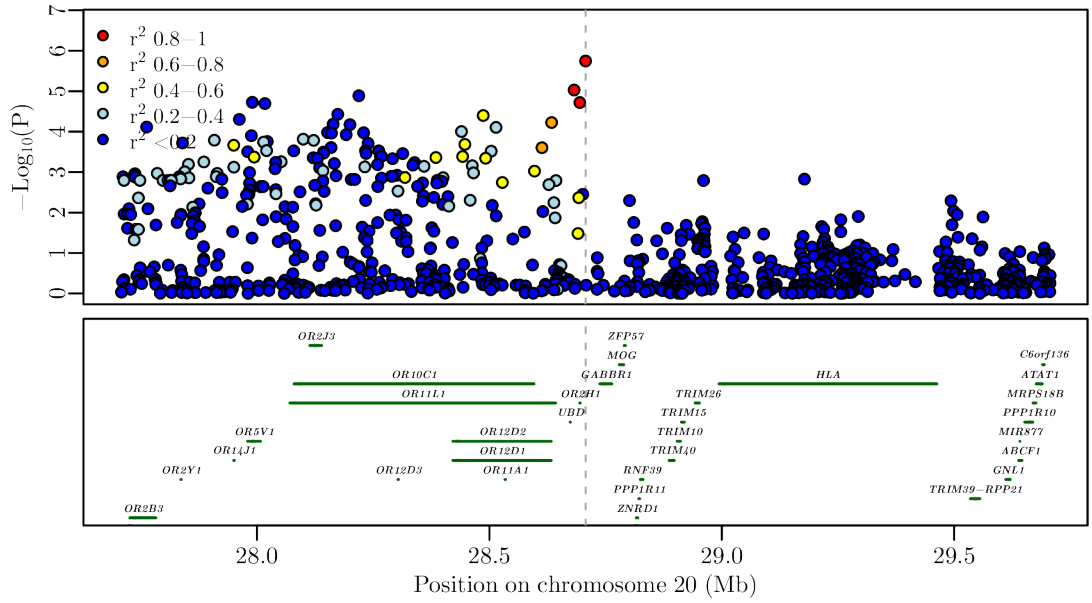
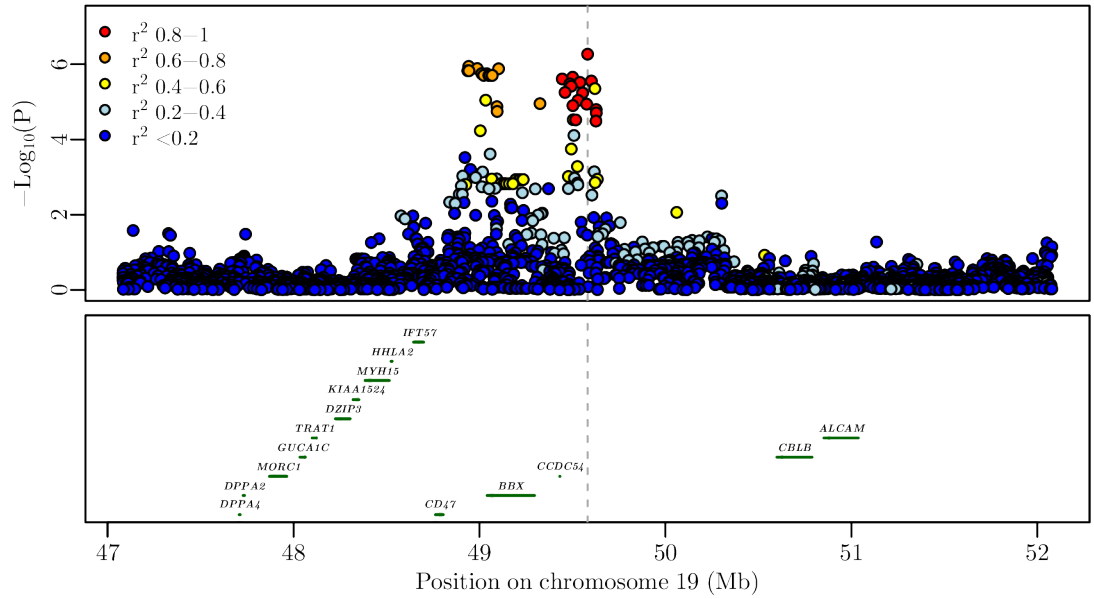


Figure D.10: Leptin loci

(a)



(b)

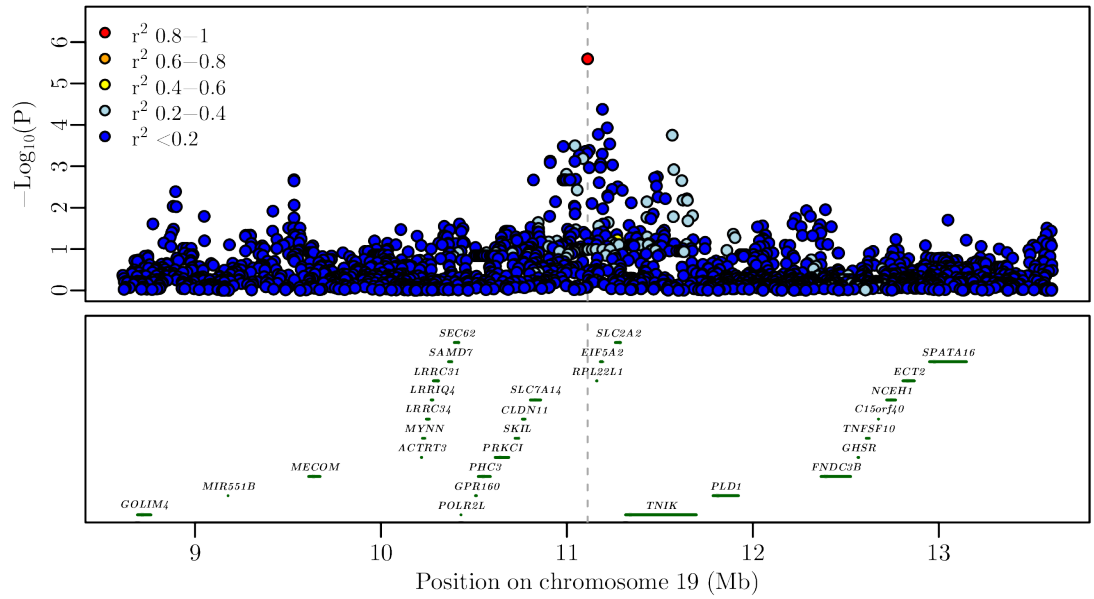
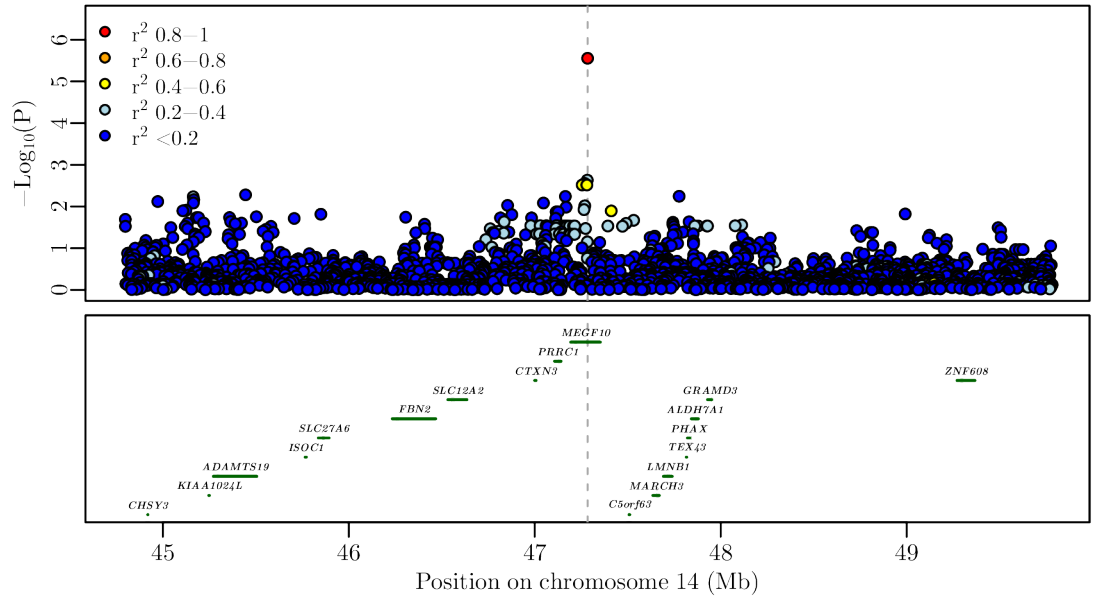


Figure D.10 *Continued from previous page*: Leptin loci

(c)



(d)

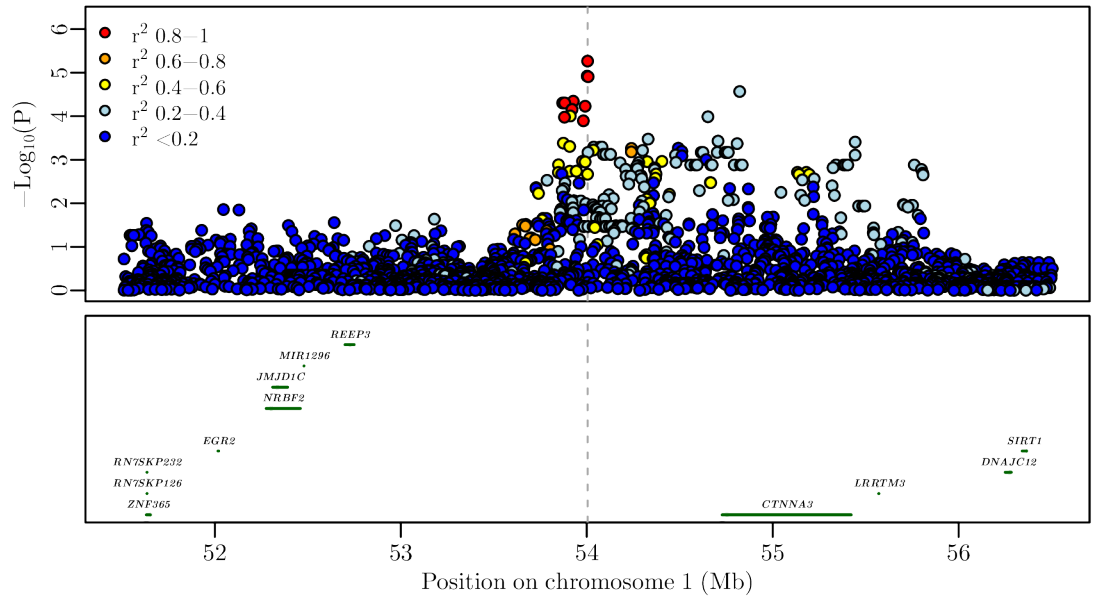
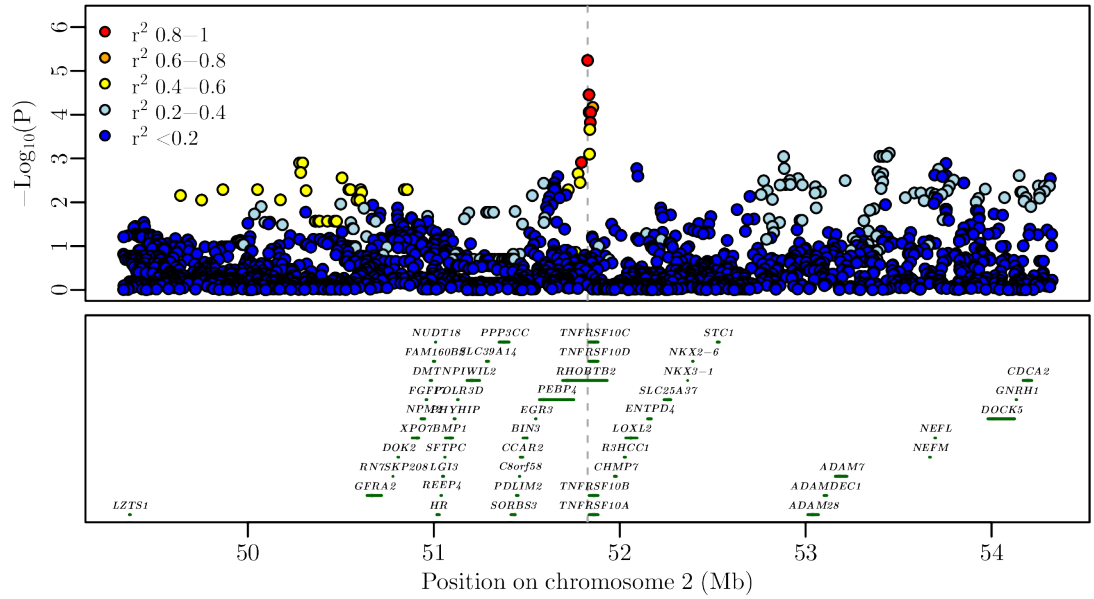


Figure D.10 *Continued from previous page*: Leptin loci

(e)



(f)

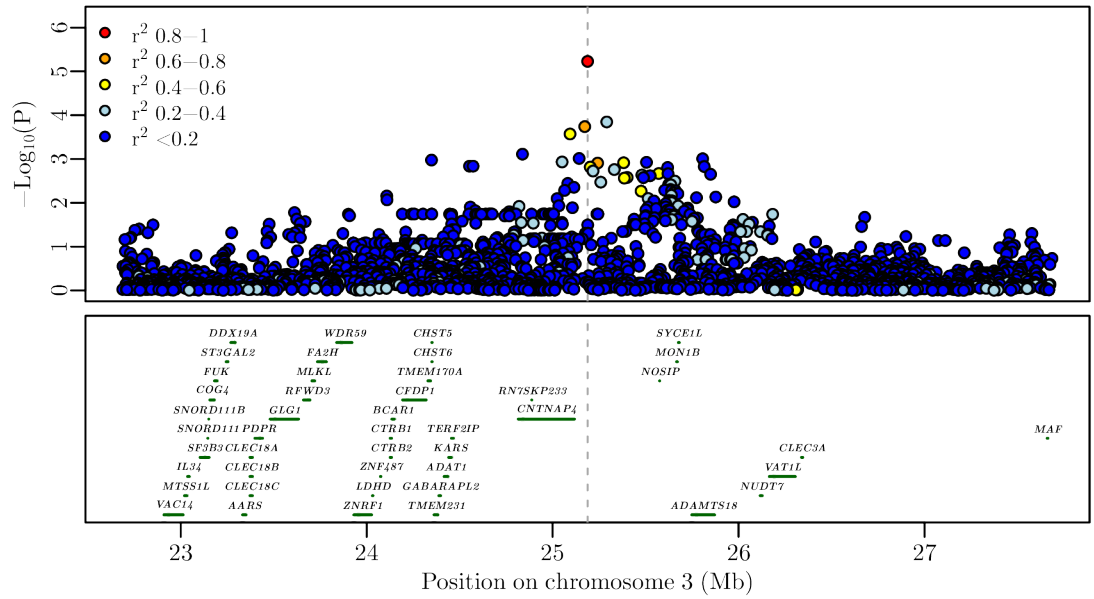


Figure D.10 *Continued from previous page*: Leptin loci

(g)

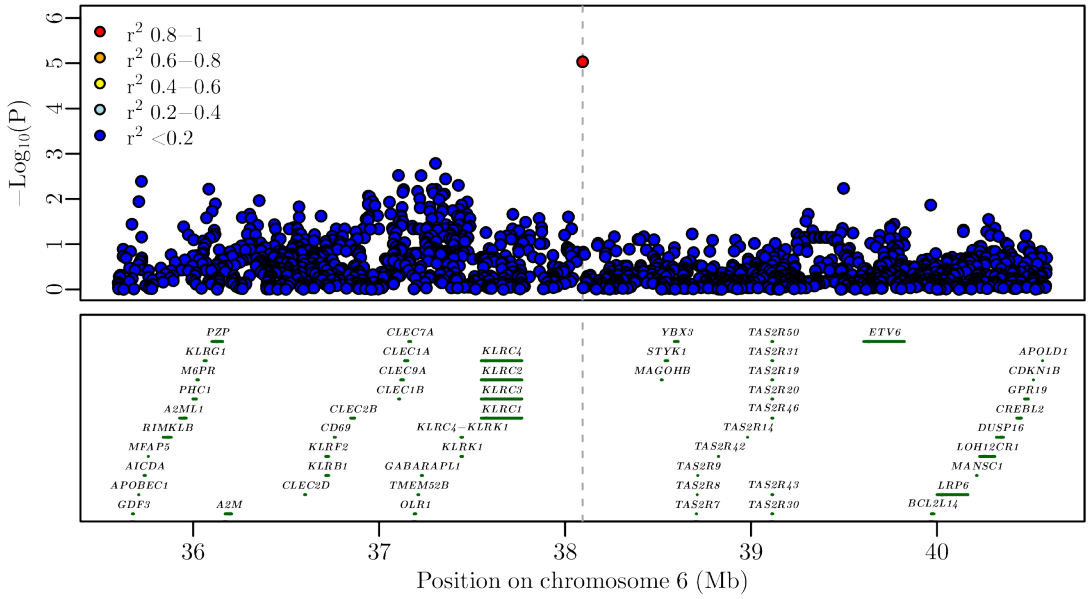


Figure D.11: Adiponectin loci

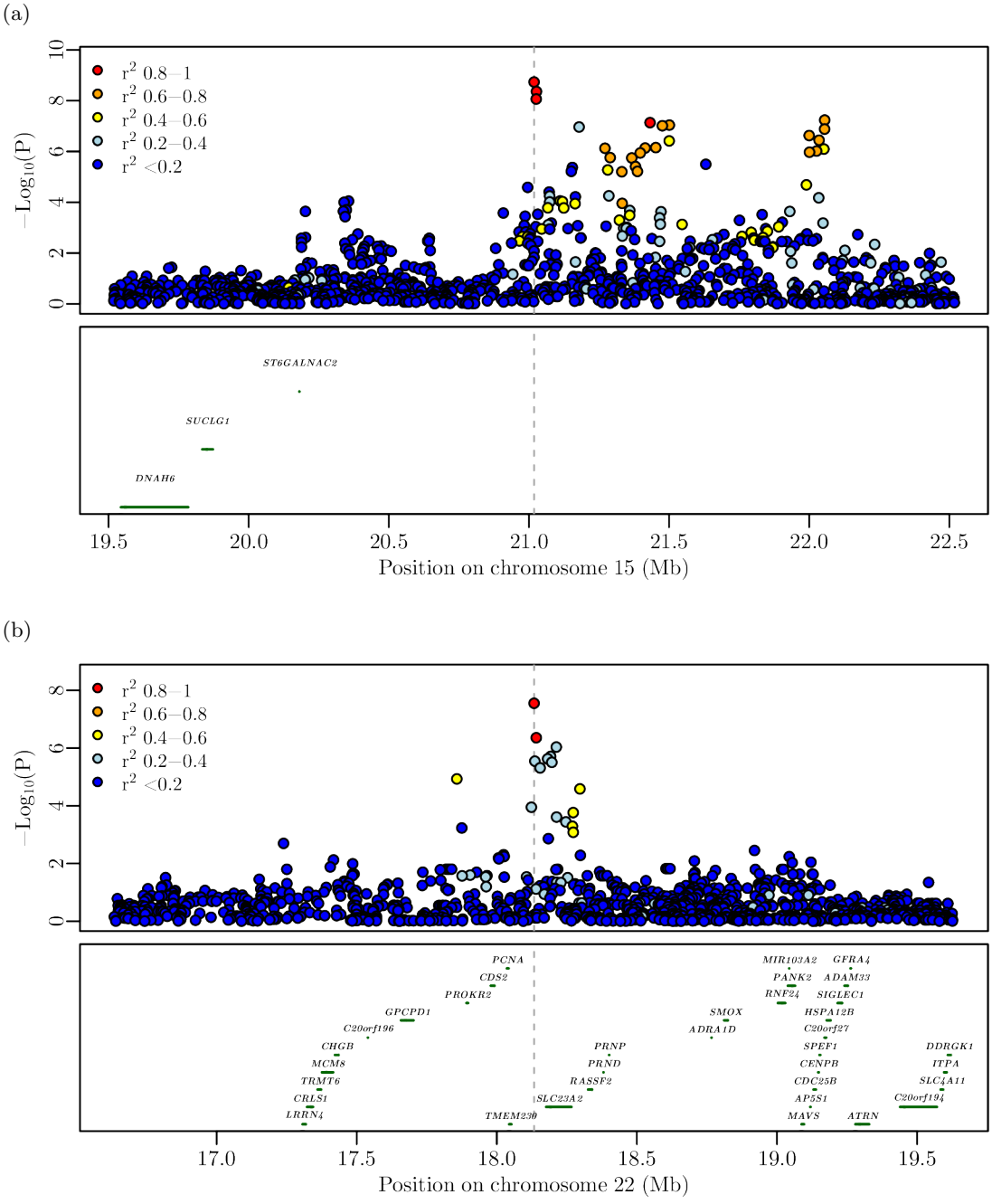
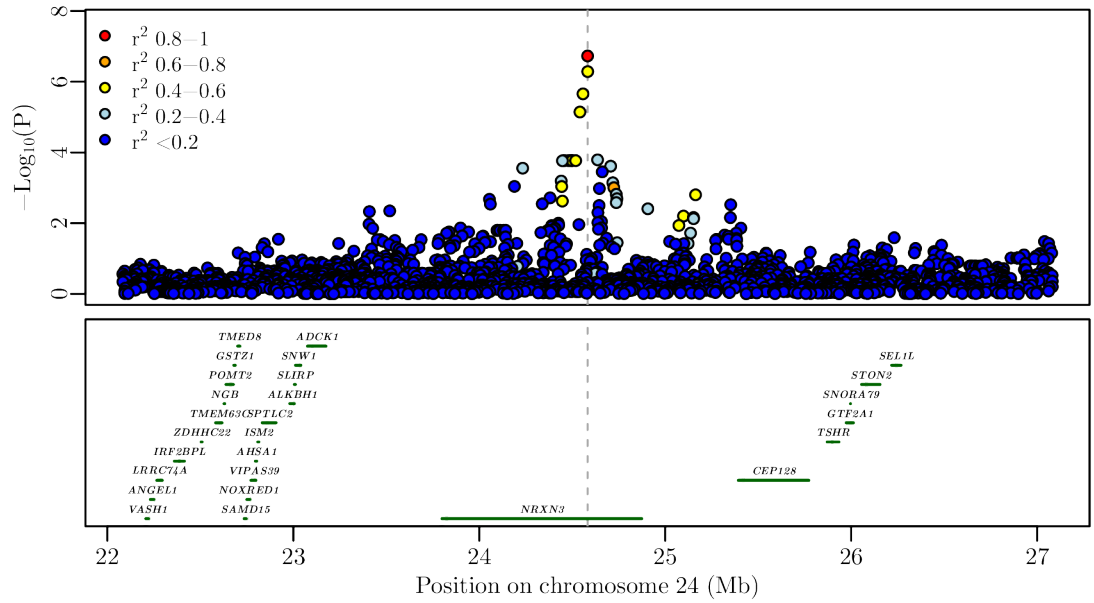


Figure D.11 *Continued from previous page*: Adiponectin loci

(c)



(d)

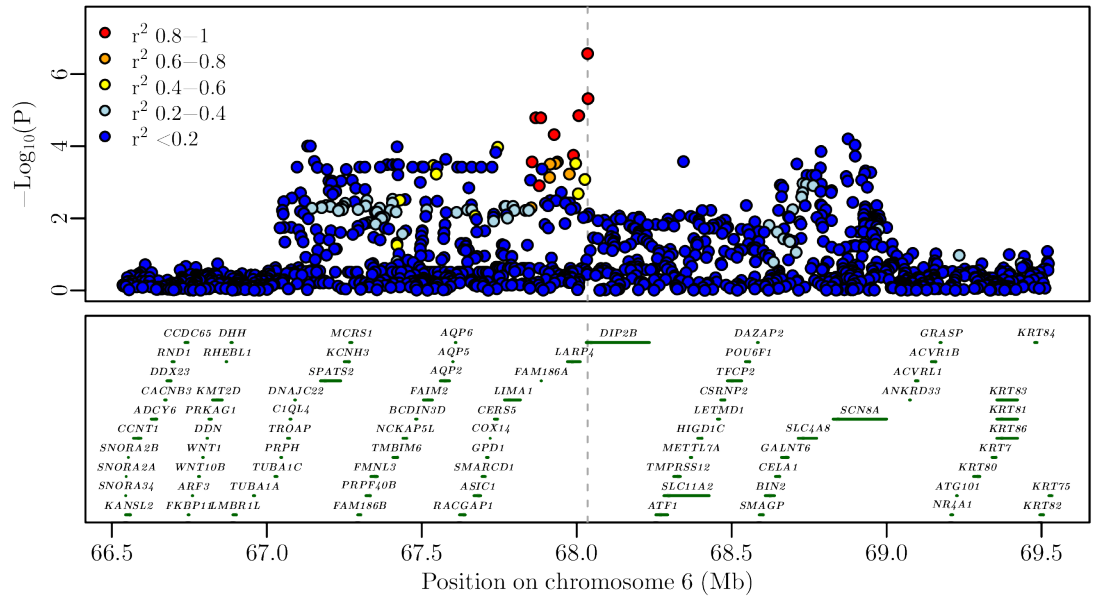
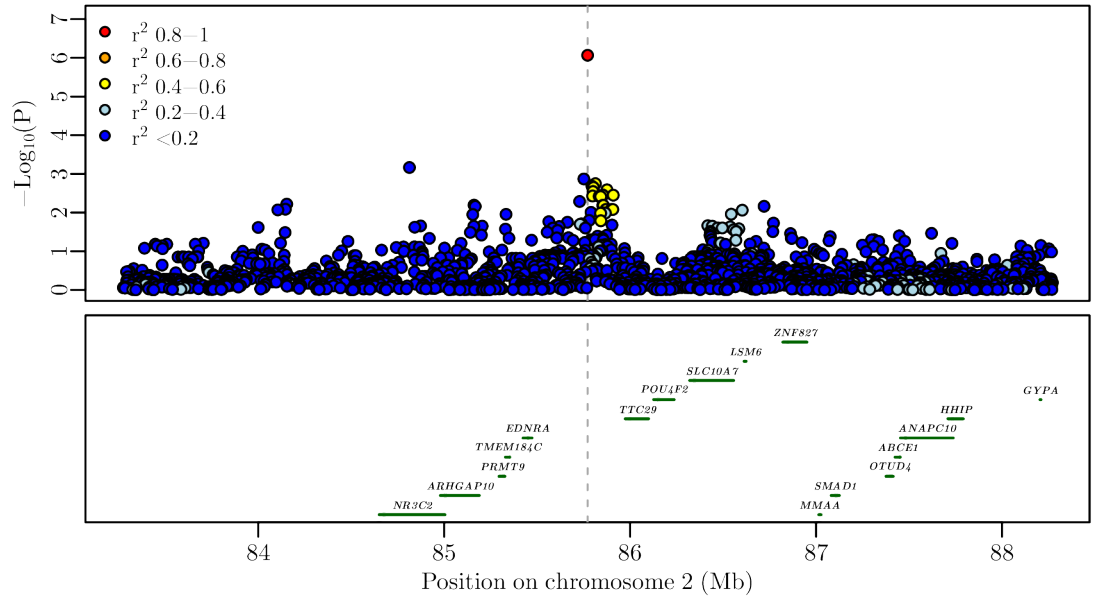


Figure D.11 *Continued from previous page*: Adiponectin loci

(e)



(f)

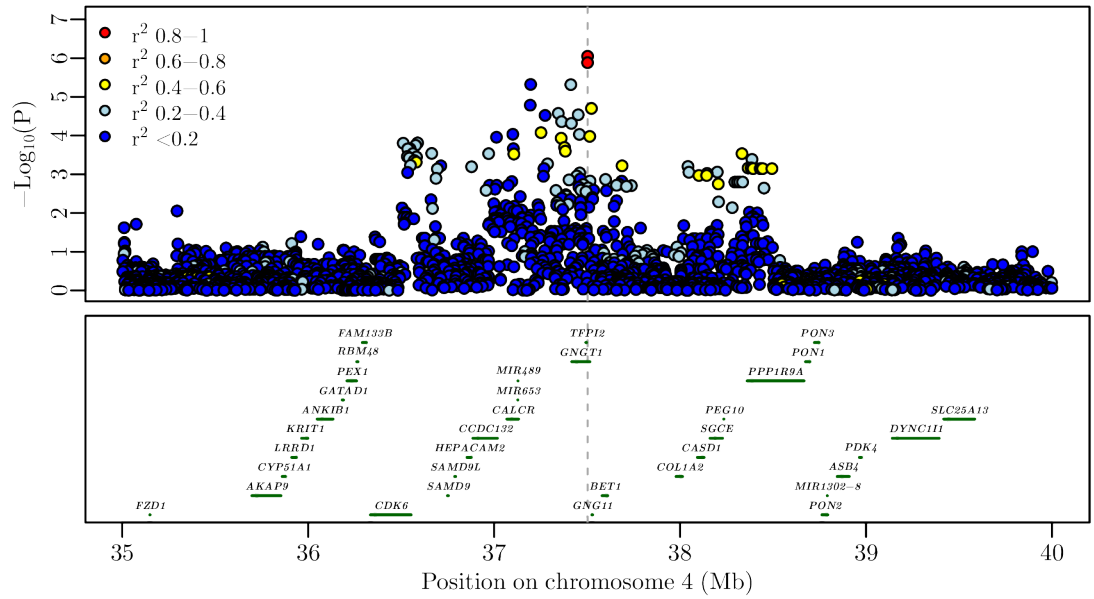
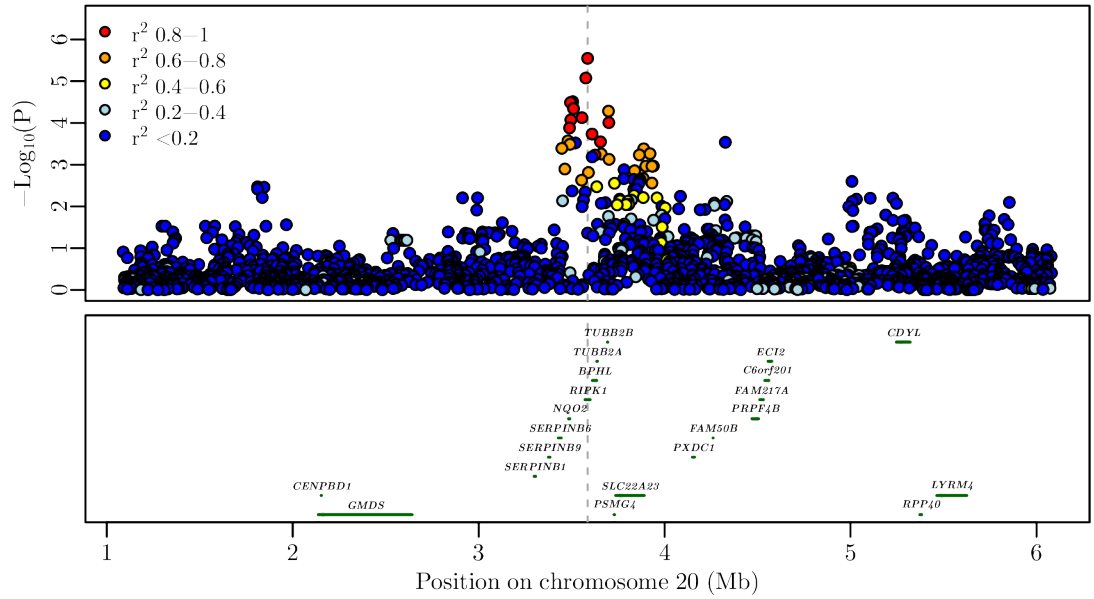


Figure D.11 *Continued from previous page*: Adiponectin loci

(g)



(h)

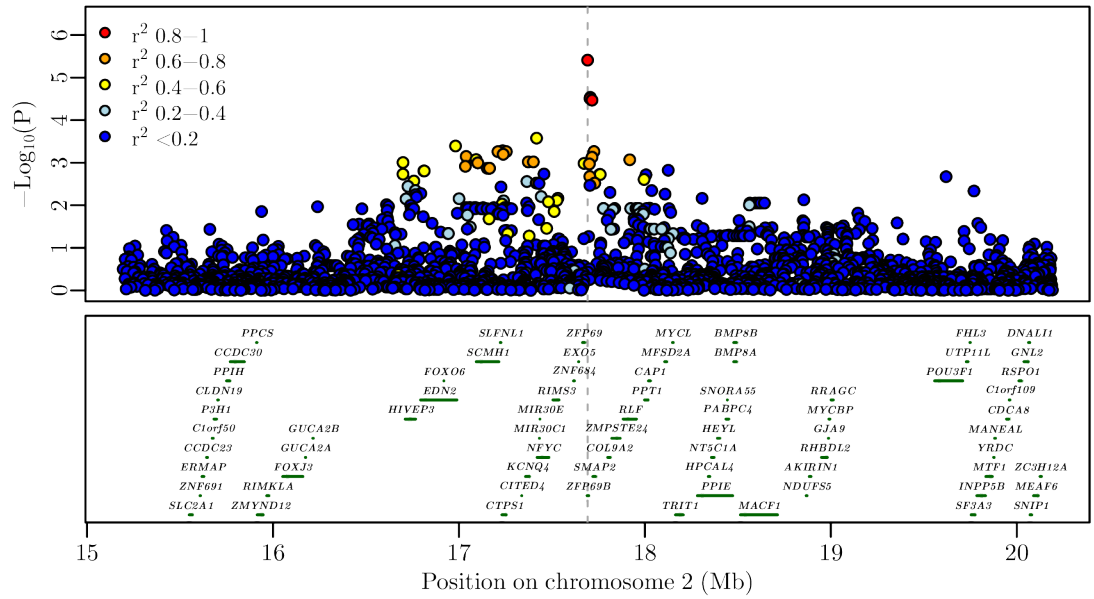
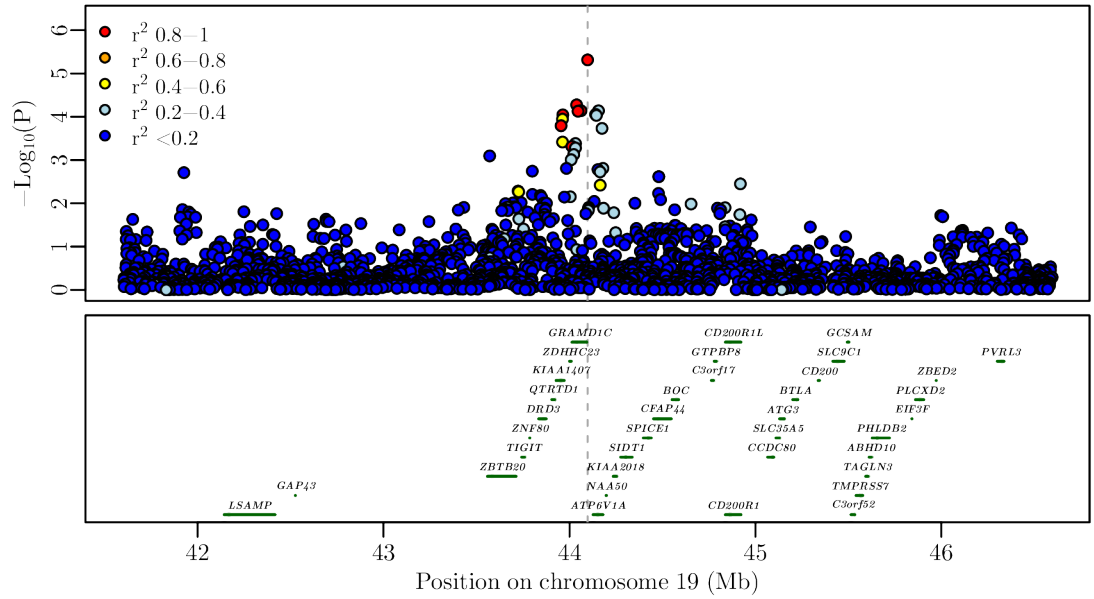


Figure D.11 *Continued from previous page*: Adiponectin loci

(i)



(j)

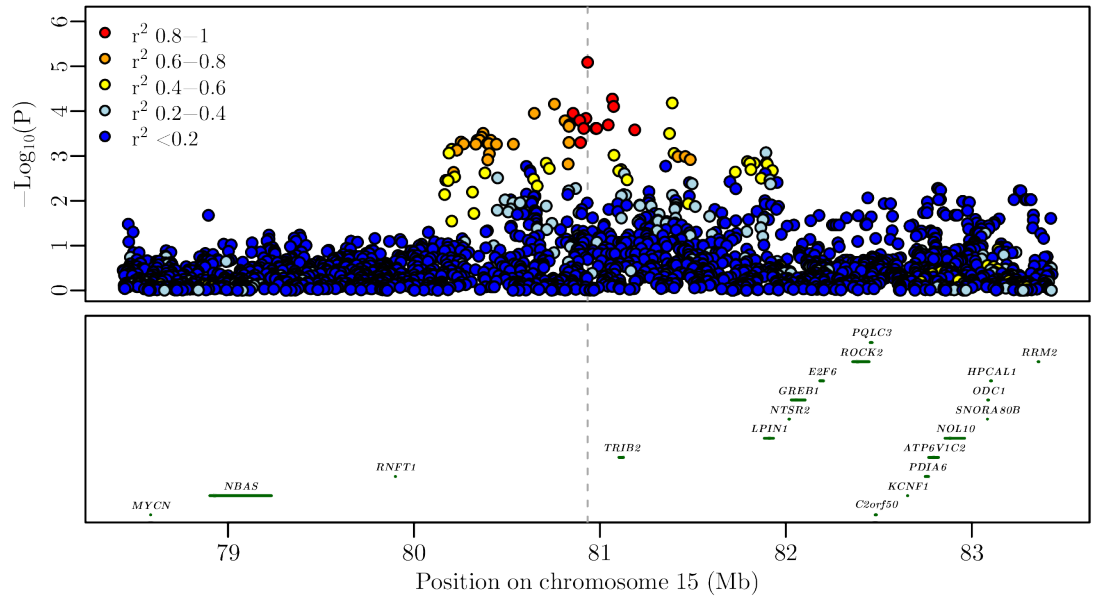


Table D.1: Significance values for all SNPs above suggestive threshold

chromosome	position	alleles	maf	beta estimate	standard error	p-value	nearest gene
17	75543887	C\T	0.3	0.21	0.04	1.86E-07	*FAM155A*
17	76068467	C\T	0.36	0.199	0.04	7.15E-07	*TNFSF13B*
17	75572753	A\G	0.3	0.201	0.041	7.53E-07	*FAM155A*
2	93491935	A\G	0.36	0.201	0.041	1.04E-06	*PCDH18*
2	93340896	G\A	0.27	0.201	0.042	1.46E-06	*PCDH18*
17	75578441	G\T	0.28	0.197	0.041	1.64E-06	*FAM155A*
4	59620539	A\G	0.08	0.2	0.042	1.96E-06	*CREB5*
19	32821180	C\T	0.23	0.198	0.042	2.53E-06	*SENP5*
4	59489648	G\A	0.09	0.194	0.041	2.60E-06	*CREB5*
2	23825094	A\C	0.08	-0.194	0.041	2.74E-06	*ZNF362*
4	59182669	T\C	0.08	0.198	0.042	2.87E-06	*CREB5*
17	76828542	T\C	0.49	0.195	0.042	2.88E-06	*MYO16*
2	93336327	G\A	0.28	0.194	0.041	2.95E-06	*PCDH18*
2	93500820	G\T	0.36	0.193	0.041	2.95E-06	*PCDH18*
4	59494006	A\G	0.09	0.192	0.041	3.04E-06	*CREB5*
17	75473022	A\G	0.42	-0.196	0.042	3.18E-06	*FAM155A*
17	75578145	C\A	0.29	0.191	0.041	3.31E-06	*FAM155A*
2	93581546	G\A	0.19	0.204	0.044	3.88E-06	*PCDH18*
19	32811045	A\C	0.23	0.193	0.042	4.10E-06	*SENP5*
4	59548966	T\C	0.15	0.192	0.042	4.15E-06	*CREB5*
4	59134336	A\C	0.08	0.194	0.042	4.26E-06	*CREB5*
4	59229403	A\C	0.08	0.194	0.042	4.26E-06	*CREB5*
4	59290113	A\G	0.08	0.194	0.042	4.26E-06	*CREB5*
1	85359274	A\C	0.09	0.192	0.042	4.85E-06	*CCSER2*
17	76846954	C\T	0.49	0.191	0.042	4.93E-06	*MYO16*
17	75618350	C\T	0.27	0.189	0.041	4.94E-06	*FAM155A*
2	93345379	A\G	0.26	0.195	0.043	5.20E-06	*PCDH18*
1	128863092	C\T	0.06	-0.18	0.039	5.23E-06	*RAB8B*
2	93390516	A\C	0.25	0.194	0.043	5.69E-06	*PCDH18*
2	93382353	T\C	0.39	0.189	0.042	5.70E-06	*PCDH18*
4	60008165	T\C	0.08	0.187	0.041	6.55E-06	*CPVL*
3	96066755	G\A	0.08	-0.194	0.043	6.66E-06	*PCDH7*
2	93488668	A\G	0.28	-0.187	0.042	6.66E-06	*PCDH18*
1	85323285	T\C	0.08	0.188	0.042	6.98E-06	*CCSER2*
19	32853596	C\T	0.23	0.188	0.042	7.15E-06	*SENP5*
2	23825921	A\C	0.08	-0.184	0.041	7.50E-06	*ZNF362*
1	85357726	C\T	0.09	0.187	0.042	7.60E-06	*CCSER2*
2	93334445	A\G	0.24	0.19	0.042	7.61E-06	*PCDH18*
2	93348200	T\C	0.24	0.19	0.042	7.61E-06	*PCDH18*
2	93363325	T\G	0.24	0.19	0.042	7.61E-06	*PCDH18*
2	93366016	T\G	0.24	0.19	0.042	7.61E-06	*PCDH18*
2	93379462	A\G	0.24	0.19	0.042	7.61E-06	*PCDH18*
2	93393451	C\A	0.24	0.19	0.042	7.61E-06	*PCDH18*
2	93407104	A\G	0.24	0.19	0.042	7.61E-06	*PCDH18*
2	93440446	G\A	0.24	0.19	0.042	7.61E-06	*PCDH18*
2	93466920	A\G	0.24	0.19	0.042	7.61E-06	*PCDH18*
2	93482049	A\G	0.24	0.19	0.042	7.61E-06	*PCDH18*
2	93524852	T\C	0.24	0.19	0.042	7.61E-06	*PCDH18*
17	76326041	C\A	0.38	0.176	0.039	7.86E-06	*MYO16*
19	32950281	G\T	0.25	0.184	0.041	7.87E-06	*PAK2*
2	93613121	A\G	0.19	0.198	0.044	7.90E-06	*PCDH18*
17	75301675	G\A	0.43	0.179	0.04	8.26E-06	*FAM155A*
4	59887211	C\T	0.09	0.185	0.042	8.63E-06	*CPVL*
4	59504852	A\G	0.08	0.188	0.042	8.65E-06	*CREB5*
4	59362148	A\C	0.1	0.187	0.042	8.66E-06	*CREB5*
2	93418906	T\C	0.26	0.189	0.042	8.78E-06	*PCDH18*

Table D.1 *Continued on next page*

Table D.1 *Continued from previous page*

chromosome	position	alleles	maf	beta estimate	standard error	p-value	nearest gene
2	93897287	A\G	0.19	0.198	0.045	8.93E-06	*HADHB*
2	93968327	T\G	0.19	0.198	0.045	8.93E-06	*HADHB*
2	93985100	G\T	0.19	0.198	0.045	8.93E-06	*HADHB*
2	93562757	T\C	0.25	0.188	0.042	9.04E-06	*PCDH18*
10	52269721	C\T	0.11	-0.189	0.043	9.32E-06	*GRIK2*
17	75656566	A\G	0.27	0.184	0.042	9.82E-06	*FAM155A*

Table D.2: Significance values for all girth:height SNPs above suggestive threshold

chromosome	position	alleles	maf	beta estimate	standard error	p-value	nearest gene
7	26154989	G\A	0.49	-0.198	0.038	1.95E-07	*MPZL2*
7	26155211	G\A	0.49	0.197	0.038	2.00E-07	*MPZL2*
7	26138800	T\C	0.43	-0.194	0.039	4.96E-07	*MPZL2*
7	26153883	C\T	0.48	-0.189	0.038	8.93E-07	*MPZL2*
7	23150348	C\T	0.07	-0.19	0.039	9.98E-07	*NXPE2*
7	22091943	C\T	0.1	-0.187	0.039	1.32E-06	*TMPRSS5*
22	10321507	A\G	0.2	0.189	0.039	1.44E-06	*SEL1L2*
7	26144328	T\C	0.44	-0.182	0.038	1.96E-06	*MPZL2*
22	9808709	C\T	0.06	0.189	0.041	2.91E-06	*FLRT3*
7	23139980	T\C	0.06	-0.18	0.039	2.98E-06	*NXPE2*
1	78254375	G\A	0.25	0.186	0.04	3.19E-06	*MAP10*
7	26148134	A\G	0.45	-0.176	0.038	3.75E-06	*MPZL2*
14	67885725	C\A	0.08	0.165	0.036	4.52E-06	*SLCO4C1*
22	11401249	T\C	0.15	0.176	0.038	4.95E-06	*ISM1*
22	11406977	T\G	0.15	0.175	0.039	5.51E-06	*ISM1*
6	50144883	A\C	0.44	-0.173	0.038	6.42E-06	*SOX5*
1	78195470	C\T	0.26	0.181	0.04	6.47E-06	*MAP10*
7	23131620	G\A	0.06	-0.175	0.039	7.25E-06	*NXPE2*
3	48758732	T\C	0.14	-0.184	0.041	7.34E-06	*GPRIN3*
20	46160950	A\G	0.28	-0.172	0.039	8.29E-06	*GPR115*
1	78578265	G\A	0.29	0.179	0.04	8.58E-06	*SIPA1L2*
22	11403577	G\A	0.15	0.173	0.039	8.70E-06	*ISM1*
1	181616615	T\C	0.38	-0.174	0.039	9.21E-06	*MDGA2*
7	26147791	C\A	0.38	0.166	0.038	9.49E-06	*MPZL2*

Table D.3: Significance values for all fasting glucose SNPs above suggestive threshold

chromosome	position	alleles	maf	beta estimate	standard error	p-value	nearest gene
8	9312611	T\C	0.31	0.24	0.046	1.55E-07	*CRYBA4*
22	26963686	A\G	0.35	0.231	0.046	5.48E-07	*DLGAP4*
22	26976287	T\C	0.35	0.228	0.046	7.04E-07	*DLGAP4*

Table D.4: Significance values for all fasting insulin SNPs above suggestive threshold

chromosome	position	alleles	maf	beta estimate	standard error	p-value	nearest gene
10	53891179	T\C	0.15	0.192	0.034	9.54E-09	*GRIK2*
24	3597984	C\T	0.06	0.193	0.034	1.02E-08	*ATG14*
24	3595302	C\T	0.07	0.19	0.034	1.57E-08	*ATG14*
24	3595016	G\A	0.09	0.186	0.034	4.62E-08	*ATG14*
24	3597697	G\A	0.07	0.18	0.035	1.92E-07	*ATG14*
19	18465605	G\A	0.19	0.173	0.033	2.27E-07	*MFN1*
19	4295255	C\T	0.36	-0.177	0.034	2.32E-07	*OTOL1*
19	4296223	C\T	0.33	-0.173	0.034	3.39E-07	*OTOL1*
1	128192334	A\G	0.25	0.167	0.033	4.13E-07	*DAPK2*
4	49751990	A\G	0.48	0.176	0.035	4.53E-07	*AHR*
24	3593543	T\C	0.08	0.172	0.034	5.41E-07	*ATG14*
1	128205038	G\A	0.26	0.163	0.033	7.02E-07	*DAPK2*
19	4295476	A\G	0.34	-0.165	0.033	8.06E-07	*OTOL1*
4	49764439	C\T	0.47	0.173	0.035	9.28E-07	*AHR*
1	128156735	T\C	0.26	0.159	0.033	1.21E-06	*DAPK2*
17	75473022	A\G	0.42	-0.163	0.034	1.35E-06	*FAM155A*
1	128228461	G\A	0.25	0.154	0.032	1.36E-06	*DAPK2*
4	49752757	C\T	0.46	0.172	0.036	1.50E-06	*AHR*
4	49755552	T\C	0.45	0.172	0.036	1.58E-06	*AHR*
4	49767207	A\G	0.47	0.167	0.035	1.68E-06	*AHR*
1	128212104	T\C	0.26	0.157	0.033	1.70E-06	*DAPK2*
20	4333850	G\A	0.35	0.168	0.036	2.41E-06	*FAM50B*
4	49765839	G\A	0.45	0.169	0.036	2.47E-06	*AHR*
12	2818242	A\G	0.21	0.164	0.035	2.94E-06	*COMMD9*
9	58656992	T\C	0.31	-0.157	0.034	4.63E-06	*TRPS1*
10	53881774	T\C	0.14	0.155	0.034	4.66E-06	*GRIK2*
2	81787272	T\C	0.43	-0.149	0.033	4.79E-06	*GATB*
4	49754549	C\T	0.46	0.167	0.037	5.20E-06	*AHR*
19	4306032	A\G	0.35	-0.153	0.034	6.52E-06	*OTOL1*
19	4308928	C\A	0.33	-0.152	0.034	7.76E-06	*OTOL1*
30	24253197	C\T	0.42	-0.15	0.034	7.78E-06	*KCNT2*
12	2358224	A\G	0.17	0.158	0.036	9.18E-06	*TRIM44*
2	81787453	A\G	0.42	-0.145	0.033	9.70E-06	*GATB*

Table D.5: Significance values for all 75 minute post-OST glucose SNPs above suggestive threshold

chromosome	position	alleles	maf	beta estimate	standard error	p-value	nearest gene
6	4059449	G\A	0.15	-0.211	0.044	1.30E-06	*SPAG16*
2	24516703	T\C	0.19	0.218	0.045	1.44E-06	*HDAC1*
2	24505077	T\C	0.19	0.214	0.045	2.43E-06	*HDAC1*
4	21624394	T\C	0.22	-0.219	0.046	2.49E-06	*COBL*
17	20363782	T\G	0.05	-0.212	0.045	2.67E-06	*DLEU7*
4	21702929	T\C	0.22	-0.214	0.047	4.41E-06	*COBL*
4	21650242	T\G	0.22	-0.212	0.047	6.06E-06	*COBL*
4	21658375	A\G	0.22	-0.212	0.047	6.06E-06	*COBL*
17	21698394	C\T	0.06	-0.202	0.045	7.26E-06	*CAB39L*
17	21713297	G\A	0.06	-0.202	0.045	7.26E-06	*CAB39L*
20	63223615	T\C	0.32	0.23	0.052	9.47E-06	*RIMS1*
26	15146523	C\T	0.08	0.198	0.045	9.84E-06	*USP25*

Table D.6: Significance values for all 75 minute post-OST insulin SNPs above suggestive threshold

chromosome	position	alleles	maf	beta estimate	standard error	p-value	nearest gene
21	20162158	C\A	0.12	-0.226	0.034	3.11E-11	*ISL1*
10	71766565	G\A	0.17	0.225	0.036	3.39E-10	*RNF217*
10	71762350	T\C	0.17	0.225	0.036	4.37E-10	*RNF217*
13	22385480	C\T	0.07	0.19	0.034	2.13E-08	*ZKSCAN2*
13	22420537	A\G	0.06	0.187	0.034	2.76E-08	*ZKSCAN2*
13	22421004	A\G	0.06	0.185	0.034	4.03E-08	*ZKSCAN2*
13	22437535	G\A	0.06	0.186	0.034	4.40E-08	*AQP8*
13	22452201	A\G	0.06	0.178	0.033	9.48E-08	*AQP8*
4	17084310	T\C	0.16	-0.175	0.033	1.16E-07	*TNS3*
10	70974737	A\C	0.09	0.184	0.035	1.59E-07	*TRDN*
10	71750321	C\T	0.21	0.186	0.036	2.26E-07	*RNF217*
4	17092940	T\G	0.15	-0.172	0.033	2.59E-07	*TNS3*
15	71121201	T\C	0.42	-0.175	0.035	3.99E-07	*NCOA1*
13	22463159	G\T	0.05	0.17	0.034	4.09E-07	*LCMT1*
10	16809864	A\C	0.34	-0.163	0.032	4.93E-07	*DACT3*
13	22419198	A\C	0.06	0.168	0.034	6.15E-07	*ZKSCAN2*
10	71796588	A\G	0.12	0.183	0.037	6.16E-07	*RNF217*
21	20164143	G\A	0.19	-0.179	0.036	6.42E-07	*ISL1*
10	71858123	C\T	0.13	0.182	0.037	9.55E-07	*RNF217*
21	20163997	T\C	0.16	-0.174	0.036	1.12E-06	*ISL1*
15	71183475	T\C	0.43	-0.168	0.035	1.21E-06	*NCOA1*
13	22369126	C\T	0.07	0.167	0.035	1.36E-06	*ZKSCAN2*
20	42730981	G\T	0.08	0.16	0.033	1.51E-06	*VEGFA*
15	71182984	C\T	0.44	-0.165	0.034	1.56E-06	*NCOA1*
10	71679132	A\G	0.11	0.177	0.037	1.84E-06	*RNF217*
4	17129514	C\T	0.24	-0.157	0.033	1.87E-06	*TNS3*
13	22444899	A\G	0.05	0.158	0.034	2.90E-06	*AQP8*
15	70830218	A\G	0.37	0.16	0.034	2.94E-06	*EFR3B*
4	17085239	C\T	0.16	-0.156	0.033	3.06E-06	*TNS3*
10	71826868	G\A	0.13	0.172	0.037	3.63E-06	*RNF217*
5	39031828	C\T	0.18	0.154	0.034	4.10E-06	*OR10K1*
4	17189808	A\G	0.16	-0.151	0.033	4.88E-06	*TNS3*
13	22430351	G\A	0.05	0.154	0.034	4.94E-06	*ZKSCAN2*
4	17182445	G\A	0.41	-0.154	0.034	5.01E-06	*TNS3*
17	1746473	C\A	0.24	0.163	0.036	5.26E-06	*XPO4*
14	65684935	A\G	0.13	-0.155	0.034	5.67E-06	*NUDT12*
15	71094746	G\A	0.36	0.155	0.034	5.99E-06	*NCOA1*
15	71174950	G\A	0.36	0.155	0.034	5.99E-06	*NCOA1*
15	71209535	G\A	0.36	0.155	0.034	5.99E-06	*NCOA1*
2	16703018	G\A	0.15	0.15	0.033	6.10E-06	*HIVEP3*
13	22512385	C\T	0.06	0.153	0.034	6.89E-06	*LCMT1*
13	22326057	A\G	0.12	0.161	0.036	7.11E-06	*ZKSCAN2*
15	71016734	A\G	0.46	-0.154	0.035	8.82E-06	*ADCY3*
15	71018294	T\C	0.46	-0.154	0.035	8.82E-06	*ADCY3*
15	71041540	G\A	0.46	-0.154	0.035	8.82E-06	*ADCY3*
21	20289132	G\T	0.16	-0.156	0.035	9.19E-06	*ISL1*
10	16768683	T\C	0.16	-0.153	0.035	9.31E-06	*DACT3*
14	65168563	A\G	0.1	-0.154	0.035	9.41E-06	*EFNA5*
4	16972268	G\T	0.12	-0.152	0.034	9.69E-06	*TNS3*
15	71059523	G\T	0.37	0.148	0.034	9.86E-06	*ADCY3*
10	71840480	A\C	0.13	0.163	0.037	9.98E-06	*RNF217*

Table D.7: Significance values for all triglyceride SNPs above suggestive threshold

chromosome	position	alleles	maf	beta estimate	standard error	p-value	nearest gene
10	64538148	T\C	0.15	0.207	0.038	7.80E-08	*FRK*
10	64250362	T\C	0.11	0.193	0.038	2.92E-07	*FRK*
9	12293620	T\C	0.18	-0.19	0.038	8.18E-07	*JPH1*
9	12289659	A\G	0.18	-0.19	0.039	8.55E-07	*JPH1*
10	64589149	G\A	0.13	0.187	0.038	9.39E-07	*FRK*
4	95568021	T\C	0.17	0.179	0.037	1.08E-06	*TRBV19*
10	64305449	T\C	0.13	0.183	0.038	1.71E-06	*FRK*
10	64391361	T\G	0.13	0.181	0.038	2.14E-06	*FRK*
15	47489715	A\G	0.4	0.181	0.038	2.29E-06	*GPR75-ASB3*
24	43212532	A\C	0.28	-0.174	0.037	2.79E-06	*DIO3*
10	64616202	T\C	0.13	0.18	0.039	2.99E-06	*FRK*
10	64257474	C\T	0.14	0.178	0.038	3.08E-06	*FRK*
10	64644659	A\G	0.15	0.177	0.038	3.44E-06	*FRK*
9	12361673	C\T	0.14	-0.179	0.039	4.04E-06	*JPH1*
10	64157652	A\G	0.17	0.174	0.038	4.35E-06	*FRK*
9	12258786	G\A	0.12	-0.17	0.037	4.44E-06	*GDAP1*
10	64295072	G\T	0.1	0.17	0.037	5.85E-06	*FRK*
2	17237220	A\G	0.08	-0.183	0.04	6.35E-06	*CTPS1*
10	64339819	G\A	0.14	0.171	0.038	6.74E-06	*FRK*
9	12360677	T\C	0.15	-0.173	0.039	7.23E-06	*JPH1*
1	15523210	C\T	0.41	0.167	0.037	7.31E-06	*PNLIPRP1*
2	17209134	A\G	0.08	-0.182	0.041	7.31E-06	*SCMH1*
2	17213809	T\C	0.08	-0.182	0.041	7.31E-06	*SCMH1*
2	17215838	G\T	0.08	-0.182	0.041	7.31E-06	*SCMH1*
10	64256124	C\T	0.12	0.172	0.038	8.04E-06	*FRK*
2	16742567	A\G	0.08	-0.179	0.04	8.83E-06	*HIVEP3*
2	16751948	T\C	0.08	-0.179	0.04	8.83E-06	*HIVEP3*
2	16756518	A\G	0.08	-0.179	0.04	8.83E-06	*HIVEP3*
2	17045837	A\C	0.08	-0.179	0.04	8.83E-06	*SCMH1*
2	17048652	T\C	0.08	-0.179	0.04	8.83E-06	*SCMH1*
2	17087976	A\G	0.08	-0.179	0.04	8.83E-06	*SCMH1*
2	17119097	T\C	0.08	-0.179	0.04	8.83E-06	*SCMH1*
2	17124058	A\G	0.08	-0.179	0.04	8.83E-06	*SCMH1*
2	17151276	A\G	0.08	-0.179	0.04	8.83E-06	*SCMH1*
2	17261734	T\G	0.08	-0.179	0.04	8.83E-06	*CTPS1*
2	17268072	C\T	0.08	-0.179	0.04	8.83E-06	*CTPS1*
2	17269106	G\T	0.08	-0.179	0.04	8.83E-06	*CTPS1*
2	17274166	A\G	0.08	-0.179	0.04	8.83E-06	*CTPS1*
2	17275705	T\G	0.08	-0.179	0.04	8.83E-06	*CTPS1*
2	17282903	G\A	0.08	-0.179	0.04	8.83E-06	*CTPS1*

Table D.8: Significance values for all NEFA SNPs above suggestive threshold

chromosome	position	alleles	maf	beta estimate	standard error	p-value	nearest gene
18	8562608	G\A	0.33	0.227	0.033	4.36E-12	*TSN*
15	66792714	G\T	0.11	0.212	0.031	8.33E-12	*ALK*
18	8545866	A\G	0.33	0.221	0.033	1.73E-11	*TSN*
15	66684047	T\G	0.1	0.208	0.031	2.06E-11	*ALK*
15	66696394	A\G	0.1	0.208	0.031	2.06E-11	*ALK*
15	66738686	G\A	0.1	0.208	0.031	2.06E-11	*ALK*
15	66742079	A\G	0.1	0.208	0.031	2.06E-11	*ALK*

Table D.8 *Continued on next page*

Table D.8 *Continued from previous page*

chromosome	position	alleles	maf	beta estimate	standard error	p-value	nearest gene
15	66744419	T\G	0.1	0.208	0.031	2.06E-11	*ALK*
15	66772783	A\C	0.1	0.208	0.031	2.06E-11	*ALK*
15	66787386	T\C	0.1	0.208	0.031	2.06E-11	*ALK*
15	66791464	G\A	0.1	0.208	0.031	2.06E-11	*ALK*
15	66806634	T\C	0.1	0.208	0.031	2.06E-11	*ALK*
15	66834278	G\A	0.1	0.208	0.031	2.06E-11	*ALK*
15	66894519	T\C	0.1	0.208	0.031	2.06E-11	*ALK*
15	66916531	G\A	0.1	0.208	0.031	2.06E-11	*ALK*
1	164702224	A\G	0.18	0.206	0.031	2.35E-11	*NOVA1*
15	66625413	C\T	0.1	0.207	0.031	2.81E-11	*ALK*
15	66954402	A\G	0.1	0.207	0.031	2.81E-11	*ALK*
15	66978254	A\G	0.1	0.207	0.031	2.81E-11	*ALK*
18	8576176	C\A	0.33	0.219	0.033	3.59E-11	*TSN*
15	66621192	A\G	0.1	0.206	0.031	3.82E-11	*ALK*
15	66619562	C\T	0.1	0.206	0.031	4.01E-11	*ALK*
15	66592272	A\G	0.1	0.204	0.031	5.40E-11	*YPEL5*
15	66604806	G\A	0.1	0.204	0.031	5.40E-11	*YPEL5*
15	66427077	A\G	0.1	0.202	0.031	6.00E-11	*LBH*
15	66853554	C\T	0.1	0.204	0.031	8.19E-11	*ALK*
18	8583757	T\C	0.36	0.215	0.033	1.01E-10	*TSN*
15	66432918	G\T	0.1	0.199	0.031	1.39E-10	*LBH*
18	8494580	G\A	0.33	0.212	0.033	1.74E-10	*TSN*
18	8498505	T\C	0.37	0.211	0.033	1.78E-10	*TSN*
24	20975408	G\T	0.35	0.2	0.031	1.83E-10	*FLVCR2*
15	66980572	T\C	0.1	0.2	0.031	1.83E-10	*ALK*
18	8510544	T\C	0.32	0.213	0.034	1.95E-10	*TSN*
18	8526404	T\C	0.33	0.209	0.033	2.44E-10	*TSN*
15	66363627	A\G	0.1	0.196	0.031	2.57E-10	*LBH*
15	66315808	T\C	0.1	0.196	0.031	2.65E-10	*LBH*
15	66364189	A\G	0.1	0.196	0.031	2.65E-10	*LBH*
15	66230520	G\A	0.11	0.196	0.031	3.57E-10	*LCLAT1*
18	8522040	G\T	0.33	0.208	0.033	4.39E-10	*TSN*
15	66360884	C\T	0.1	0.193	0.031	4.56E-10	*LBH*
15	66257969	C\T	0.1	0.193	0.031	5.65E-10	*LCLAT1*
15	66263349	T\C	0.1	0.193	0.031	5.65E-10	*LCLAT1*
15	66281832	G\A	0.1	0.193	0.031	5.65E-10	*LCLAT1*
15	66312187	C\T	0.1	0.193	0.031	5.65E-10	*LBH*
18	8552881	A\G	0.37	0.202	0.033	6.20E-10	*TSN*
15	66353731	G\A	0.1	0.192	0.031	6.25E-10	*LBH*
15	66354719	C\T	0.1	0.192	0.031	6.25E-10	*LBH*
15	66362587	T\C	0.1	0.192	0.031	6.25E-10	*LBH*
15	66368420	T\G	0.1	0.192	0.031	6.25E-10	*LBH*
15	66370910	C\A	0.1	0.192	0.031	6.25E-10	*LBH*
15	66897447	A\G	0.06	0.199	0.032	9.44E-10	*ALK*
1	183728003	G\A	0.23	-0.192	0.031	9.68E-10	*KLHDC1*
18	8541982	A\G	0.37	0.204	0.033	9.88E-10	*TSN*
18	8518139	G\A	0.4	0.196	0.032	1.06E-09	*TSN*
18	8868366	C\T	0.31	0.205	0.034	1.07E-09	*TSN*
15	66271610	A\G	0.1	0.188	0.031	1.27E-09	*LCLAT1*
18	8898085	A\G	0.3	0.205	0.034	1.44E-09	*TSN*
15	66451849	A\G	0.07	0.194	0.032	1.53E-09	*LBH*
15	66203067	C\A	0.12	0.188	0.031	2.23E-09	*LCLAT1*
15	66189967	C\T	0.12	0.185	0.031	3.26E-09	*LCLAT1*
15	67012184	T\C	0.1	0.187	0.032	3.26E-09	*ALK*
15	67056101	T\C	0.1	0.187	0.032	3.26E-09	*ALK*
15	67061889	G\A	0.1	0.187	0.032	3.26E-09	*ALK*
15	67081698	C\T	0.1	0.187	0.032	3.26E-09	*ALK*

Table D.8 *Continued on next page*

Table D.8 *Continued from previous page*

chromosome	position	alleles	maf	beta estimate	standard error	p-value	nearest gene
15	67107022	T\C	0.1	0.187	0.032	3.26E-09	*ALK*
15	67122584	A\G	0.1	0.187	0.032	3.26E-09	*ALK*
15	67125854	A\C	0.1	0.187	0.032	3.26E-09	*ALK*
15	67155661	C\A	0.1	0.187	0.032	3.26E-09	*ALK*
15	67171958	T\G	0.1	0.187	0.032	3.26E-09	*ALK*
15	67183880	T\C	0.1	0.187	0.032	3.26E-09	*ALK*
15	67213022	T\C	0.1	0.187	0.032	3.26E-09	*ALK*
15	67213270	T\C	0.1	0.187	0.032	3.26E-09	*ALK*
15	67228685	A\C	0.1	0.187	0.032	3.26E-09	*ALK*
15	67229504	A\G	0.1	0.187	0.032	3.26E-09	*ALK*
15	67236481	A\G	0.1	0.187	0.032	3.26E-09	*ALK*
15	67260893	T\C	0.1	0.187	0.032	3.26E-09	*ALK*
15	67300229	C\T	0.1	0.187	0.032	3.26E-09	*ALK*
15	67326188	A\G	0.1	0.187	0.032	3.26E-09	*ALK*
15	66184961	A\C	0.13	0.184	0.031	3.55E-09	*LCLAT1*
18	8948638	G\A	0.26	0.204	0.035	4.66E-09	*TSN*
18	9016362	G\A	0.26	0.204	0.035	4.91E-09	*NIFK*
15	66988869	A\G	0.14	0.182	0.031	5.34E-09	*ALK*
15	66141154	A\G	0.13	0.181	0.031	7.92E-09	*LCLAT1*
18	8569273	G\A	0.41	0.191	0.033	8.22E-09	*TSN*
18	8607534	T\G	0.37	0.184	0.032	9.77E-09	*TSN*
15	66110588	G\A	0.13	0.178	0.031	1.00E-08	*LCLAT1*
7	84765240	T\C	0.46	0.18	0.032	1.16E-08	*SOX6*
18	8917708	C\A	0.26	0.197	0.035	1.56E-08	*TSN*
15	66029737	G\A	0.16	0.172	0.031	1.81E-08	*CAPN13*
18	9401773	T\C	0.27	0.193	0.034	1.92E-08	*TFCP2L1*
18	8981003	A\G	0.26	0.194	0.035	2.20E-08	*NIFK*
18	8593141	T\G	0.43	0.183	0.033	2.47E-08	*TSN*
18	8657735	C\T	0.35	0.186	0.033	2.54E-08	*TSN*
15	66053380	T\C	0.14	0.176	0.032	2.61E-08	*LCLAT1*
18	8408544	A\G	0.29	0.186	0.033	2.63E-08	*TSN*
15	66032805	A\G	0.14	0.172	0.031	2.65E-08	*CAPN13*
18	8547549	C\T	0.48	0.179	0.032	2.70E-08	*TSN*
18	8649138	A\G	0.34	0.186	0.033	2.72E-08	*TSN*
1	183663863	G\T	0.19	-0.179	0.032	2.88E-08	*DNAAF2*
1	183668205	T\C	0.19	-0.179	0.032	2.88E-08	*DNAAF2*
15	66135520	T\G	0.13	0.173	0.031	3.10E-08	*LCLAT1*
18	8505327	C\T	0.42	0.185	0.033	3.41E-08	*TSN*
7	84803850	G\A	0.48	0.174	0.032	3.64E-08	*SOX6*
2	105684803	G\T	0.11	0.165	0.03	4.03E-08	*ADAD1*
18	8729227	C\A	0.45	0.174	0.032	4.06E-08	*TSN*
18	7731541	C\A	0.35	0.176	0.032	4.30E-08	*CNTNAP5*
2	105698972	A\G	0.1	0.165	0.03	4.44E-08	*ADAD1*
15	66052656	T\C	0.13	0.173	0.032	4.48E-08	*LCLAT1*
15	66061042	G\T	0.13	0.173	0.032	4.48E-08	*LCLAT1*
18	8678717	C\A	0.34	0.183	0.033	4.62E-08	*TSN*
18	8625323	T\C	0.35	0.183	0.034	4.66E-08	*TSN*
15	66948962	C\T	0.14	0.171	0.031	4.90E-08	*ALK*
15	66083987	G\A	0.13	0.17	0.031	5.06E-08	*LCLAT1*
15	66605848	A\C	0.14	0.172	0.032	5.24E-08	*YPEL5*
18	8964561	G\A	0.27	0.189	0.035	5.45E-08	*TSN*
18	8651943	T\C	0.34	0.183	0.034	5.46E-08	*TSN*
1	184178932	A\G	0.19	-0.173	0.032	5.46E-08	*ATP5S*
18	7732359	T\C	0.35	0.175	0.032	5.55E-08	*CNTNAP5*
18	8638988	C\A	0.34	0.182	0.033	5.84E-08	*TSN*
18	8407744	G\T	0.28	0.181	0.033	5.99E-08	*TSN*
15	66790529	G\A	0.15	0.172	0.032	6.42E-08	*ALK*

Table D.8 *Continued on next page*

Table D.8 *Continued from previous page*

chromosome	position	alleles	maf	beta estimate	standard error	p-value	nearest gene
1	183675692	C\T	0.19	-0.175	0.032	6.76E-08	*POLE2*
18	8699152	A\G	0.36	0.181	0.033	6.80E-08	*TSN*
1	183685920	A\C	0.19	-0.174	0.032	7.11E-08	*POLE2*
1	183729679	G\A	0.19	-0.174	0.032	7.11E-08	*KLHDC1*
1	183745396	A\G	0.19	-0.174	0.032	7.11E-08	*KLHDC1*
1	183791455	C\T	0.19	-0.174	0.032	7.11E-08	*NEMF*
18	8544905	A\G	0.48	0.172	0.032	7.69E-08	*TSN*
10	13860272	G\A	0.33	0.163	0.03	7.87E-08	*LYPD4*
15	66065116	T\C	0.13	0.169	0.031	8.03E-08	*LCLAT1*
15	66035187	T\C	0.13	0.168	0.031	8.42E-08	*LCLAT1*
15	66925052	A\G	0.14	0.169	0.031	8.49E-08	*ALK*
15	66936323	A\G	0.14	0.169	0.031	8.49E-08	*ALK*
15	66947690	A\G	0.14	0.169	0.031	8.49E-08	*ALK*
18	9410752	A\C	0.26	0.187	0.035	9.08E-08	*TFCP2L1*
7	84804729	G\A	0.46	0.168	0.031	9.12E-08	*SOX6*
18	9407363	C\T	0.31	0.182	0.034	9.27E-08	*TFCP2L1*
15	66166383	G\A	0.13	0.169	0.032	9.41E-08	*LCLAT1*
18	9040798	G\A	0.33	0.18	0.034	9.56E-08	*CLASP1*
15	66138051	C\T	0.13	0.168	0.032	1.04E-07	*LCLAT1*
18	8651718	C\T	0.34	0.179	0.034	1.05E-07	*TSN*
18	8440080	A\G	0.24	0.178	0.034	1.27E-07	*TSN*
18	8409473	T\C	0.28	0.176	0.033	1.38E-07	*TSN*
18	8997736	C\T	0.34	0.172	0.033	1.49E-07	*NIFK*
18	7863587	T\G	0.33	0.171	0.033	1.63E-07	*TSN*
2	105681701	C\A	0.1	0.159	0.03	1.71E-07	*ADAD1*
18	9090283	G\A	0.31	0.177	0.034	1.73E-07	*CLASP1*
15	66151620	T\C	0.13	0.165	0.032	1.80E-07	*LCLAT1*
18	7806575	T\C	0.34	0.169	0.033	1.96E-07	*CNTNAP5*
18	8388383	A\C	0.24	0.176	0.034	2.10E-07	*TSN*
18	7826885	T\G	0.33	0.168	0.032	2.16E-07	*TSN*
18	7827332	C\T	0.33	0.167	0.032	2.60E-07	*TSN*
1	183865146	C\T	0.19	-0.167	0.032	2.64E-07	*ARF6*
15	66120688	G\T	0.13	0.161	0.031	2.64E-07	*LCLAT1*
18	9407244	T\C	0.27	0.177	0.034	2.86E-07	*TFCP2L1*
18	7757777	T\C	0.33	0.166	0.032	2.96E-07	*CNTNAP5*
1	183813125	G\A	0.2	-0.166	0.032	2.99E-07	*NEMF*
18	7824472	G\A	0.33	0.166	0.032	3.11E-07	*TSN*
18	7758382	C\A	0.35	0.168	0.033	3.20E-07	*CNTNAP5*
18	7896646	A\G	0.33	0.167	0.033	3.36E-07	*TSN*
1	183724896	C\A	0.23	-0.164	0.032	3.39E-07	*KLHDC1*
2	105683559	A\G	0.13	0.162	0.032	3.41E-07	*ADAD1*
1	183813707	C\T	0.32	-0.161	0.032	3.47E-07	*NEMF*
18	7770342	T\C	0.33	0.165	0.032	3.56E-07	*CNTNAP5*
18	8443887	A\G	0.35	0.172	0.034	3.59E-07	*TSN*
18	7733248	T\G	0.24	0.167	0.033	3.62E-07	*CNTNAP5*
18	9031138	A\G	0.25	0.175	0.034	3.73E-07	*CLASP1*
18	8833594	G\A	0.3	0.17	0.034	3.79E-07	*TSN*
1	183829350	C\T	0.2	-0.164	0.032	3.83E-07	*NEMF*
1	183840941	C\A	0.2	-0.164	0.032	3.83E-07	*NEMF*
18	9126396	G\A	0.29	0.175	0.034	3.96E-07	*CLASP1*
15	65944235	G\A	0.11	0.158	0.031	4.03E-07	*CAPN13*
15	65776991	G\A	0.09	0.159	0.031	4.21E-07	*GALNT14*
2	1779499	T\C	0.08	0.155	0.031	4.75E-07	*OMA1*
2	105711808	C\T	0.1	0.154	0.031	5.46E-07	*ADAD1*
24	21008374	T\C	0.14	-0.159	0.032	6.14E-07	*FLVCR2*
18	9079802	G\A	0.28	0.175	0.035	6.30E-07	*CLASP1*
18	8535293	C\T	0.46	0.162	0.033	6.59E-07	*TSN*

Table D.8 *Continued on next page*

Table D.8 *Continued from previous page*

chromosome	position	alleles	maf	beta estimate	standard error	p-value	nearest gene
15	66022795	G\A	0.15	0.157	0.032	6.63E-07	*CAPN13*
24	21529408	G\A	0.22	-0.16	0.032	6.66E-07	*TGFB3*
18	8834248	T\G	0.34	0.165	0.033	6.83E-07	*TSN*
24	20468564	A\C	0.45	-0.16	0.032	7.40E-07	*EIF2B2*
18	8701876	G\A	0.5	0.164	0.033	7.41E-07	*TSN*
18	8538709	T\C	0.46	0.162	0.033	7.51E-07	*TSN*
18	8542644	G\A	0.46	0.162	0.033	7.51E-07	*TSN*
18	8729817	G\A	0.43	0.158	0.032	8.19E-07	*TSN*
10	13465934	A\G	0.45	0.156	0.032	8.57E-07	*TMEM145*
18	8863197	G\A	0.32	0.161	0.033	8.91E-07	*TSN*
18	9076539	C\T	0.28	0.172	0.035	9.17E-07	*CLASP1*
1	164921944	C\T	0.07	-0.154	0.031	9.18E-07	*NOVA1*
15	65931933	T\C	0.11	0.151	0.031	9.42E-07	*CAPN13*
18	9292376	G\A	0.31	0.167	0.034	9.77E-07	*CLASP1*
18	9102040	C\A	0.34	0.171	0.035	9.79E-07	*CLASP1*
2	105718349	T\C	0.11	0.148	0.03	1.15E-06	*ADAD1*
24	21015302	C\T	0.29	-0.149	0.031	1.21E-06	*FLVCR2*
2	105672114	A\C	0.13	0.153	0.032	1.28E-06	*ADAD1*
1	164883297	A\G	0.08	-0.153	0.032	1.31E-06	*NOVA1*
18	8547924	G\A	0.48	-0.159	0.033	1.31E-06	*TSN*
10	13850810	T\C	0.33	0.151	0.031	1.57E-06	*LYPD4*
15	65879226	C\T	0.11	0.152	0.032	1.57E-06	*GALNT14*
18	9408052	C\A	0.33	0.161	0.034	1.70E-06	*TFCP2L1*
30	19352250	T\G	0.09	0.145	0.031	1.98E-06	*BRINP3*
15	65885117	T\C	0.11	0.149	0.031	2.01E-06	*GALNT14*
24	21392431	A\G	0.15	-0.154	0.032	2.04E-06	*TTLL5*
1	183129885	C\T	0.29	0.149	0.031	2.13E-06	*RPS29*
18	8645466	T\C	0.47	0.157	0.033	2.16E-06	*TSN*
18	9039740	C\A	0.41	0.153	0.032	2.25E-06	*CLASP1*
18	8735008	C\T	0.44	0.153	0.032	2.30E-06	*TSN*
18	9306019	C\A	0.31	0.162	0.034	2.34E-06	*CLASP1*
18	8683786	A\G	0.4	0.155	0.033	2.58E-06	*TSN*
1	183808579	C\T	0.28	-0.149	0.032	2.69E-06	*NEMF*
24	21177322	A\G	0.26	-0.147	0.031	2.75E-06	*FLVCR2*
1	164735535	A\G	0.13	0.151	0.032	2.96E-06	*NOVA1*
1	183612813	A\G	0.09	-0.146	0.031	2.97E-06	*RPS29*
18	9245778	G\A	0.32	0.163	0.035	2.99E-06	*CLASP1*
1	183691527	G\A	0.28	-0.144	0.031	3.06E-06	*POLE2*
18	7745107	T\G	0.23	0.154	0.033	3.07E-06	*CNTNAP5*
18	9041020	C\A	0.34	0.161	0.035	3.12E-06	*CLASP1*
18	9063071	T\C	0.27	0.161	0.035	3.23E-06	*CLASP1*
1	164672635	C\T	0.21	-0.142	0.03	3.27E-06	*NOVA1*
1	164604740	C\T	0.15	0.149	0.032	3.29E-06	*NOVA1*
18	9152025	C\T	0.32	0.16	0.035	3.38E-06	*CLASP1*
9	74443615	C\T	0.22	-0.139	0.03	3.40E-06	*ST3GAL1*
30	20138883	G\A	0.07	0.139	0.03	3.45E-06	*BRINP3*
11	56669565	A\G	0.21	-0.145	0.031	3.50E-06	*HS3ST3B1*
1	164741882	G\T	0.12	0.147	0.032	3.51E-06	*NOVA1*
14	80043766	C\T	0.09	0.146	0.032	3.54E-06	*RN7SKP34*
14	80059792	G\A	0.09	0.146	0.032	3.54E-06	*RN7SKP34*
14	80079214	G\A	0.09	0.146	0.032	3.54E-06	*RN7SKP34*
14	80083457	T\G	0.09	0.146	0.032	3.54E-06	*RN7SKP34*
14	80121164	G\A	0.09	0.146	0.032	3.54E-06	*RN7SKP34*
14	80143894	T\G	0.09	0.146	0.032	3.54E-06	*RN7SKP34*
14	80157090	A\G	0.09	0.146	0.032	3.54E-06	*RN7SKP34*
14	80169308	T\G	0.09	0.146	0.032	3.54E-06	*RN7SKP34*
15	66307107	T\C	0.2	0.148	0.032	3.57E-06	*LCLAT1*

Table D.8 *Continued on next page*

Table D.8 *Continued from previous page*

chromosome	position	alleles	maf	beta estimate	standard error	p-value	nearest gene
1	183773554	A\G	0.28	-0.145	0.031	3.60E-06	*KLHDC2*
18	8695239	T\C	0.45	0.156	0.034	3.67E-06	*TSN*
10	14225679	A\G	0.42	0.141	0.03	3.73E-06	*CD177*
7	85044894	T\C	0.12	0.138	0.03	3.78E-06	*C11orf58*
18	7754632	G\T	0.21	0.15	0.032	3.79E-06	*CNTNAP5*
1	164863262	G\A	0.06	-0.145	0.031	3.87E-06	*NOVA1*
1	164879729	A\G	0.07	-0.146	0.032	4.08E-06	*NOVA1*
15	66891533	A\G	0.17	0.154	0.033	4.08E-06	*ALK*
9	74555200	C\T	0.11	-0.137	0.03	4.12E-06	*ZFAT*
15	66810086	C\T	0.17	0.153	0.033	4.21E-06	*ALK*
24	19986094	A\G	0.1	-0.151	0.033	4.25E-06	*ISCA2*
18	9227305	T\C	0.31	0.159	0.035	4.51E-06	*CLASP1*
14	80095754	T\C	0.1	0.144	0.032	4.72E-06	*RN7SKP34*
1	164686433	T\C	0.15	0.146	0.032	4.86E-06	*NOVA1*
18	9327923	C\T	0.32	0.157	0.034	4.88E-06	*CLASP1*
1	183829193	C\T	0.28	-0.145	0.032	5.03E-06	*NEMF*
24	20660972	T\C	0.38	-0.147	0.032	5.08E-06	*FOS*
11	56211286	G\A	0.4	-0.135	0.03	5.15E-06	*COX10*
24	21593980	T\G	0.18	-0.144	0.032	5.16E-06	*IFT43*
18	8601859	A\G	0.43	-0.156	0.034	5.26E-06	*TSN*
1	164942346	T\C	0.05	-0.142	0.031	5.30E-06	*NOVA1*
1	35305763	C\T	0.29	0.144	0.032	5.46E-06	*LGI1*
1	164712972	T\C	0.14	0.144	0.032	5.67E-06	*NOVA1*
18	8602353	C\T	0.33	-0.15	0.033	5.78E-06	*TSN*
1	184259015	G\T	0.08	-0.143	0.032	5.96E-06	*MAP4K5*
1	164648967	T\C	0.12	0.148	0.033	6.11E-06	*NOVA1*
11	56150165	A\G	0.35	-0.135	0.03	6.25E-06	*COX10*
11	56176479	G\A	0.35	-0.135	0.03	6.25E-06	*COX10*
11	56177741	G\A	0.35	-0.135	0.03	6.25E-06	*COX10*
11	56188490	T\C	0.35	-0.135	0.03	6.25E-06	*COX10*
11	56192007	G\A	0.35	-0.135	0.03	6.25E-06	*COX10*
1	164668494	A\C	0.11	0.144	0.032	6.30E-06	*NOVA1*
1	164764475	C\A	0.11	0.146	0.032	6.42E-06	*NOVA1*
1	164698836	C\T	0.12	0.143	0.032	6.45E-06	*NOVA1*
24	21010123	C\T	0.21	0.14	0.031	6.46E-06	*FLVCR2*
18	75270775	G\A	0.14	-0.136	0.03	6.47E-06	*KCTD18*
1	164953244	C\T	0.06	-0.142	0.032	6.51E-06	*NOVA1*
18	8600593	G\A	0.47	-0.154	0.034	6.80E-06	*TSN*
1	164717730	T\C	0.23	-0.139	0.031	7.09E-06	*NOVA1*
1	183910385	T\C	0.09	-0.141	0.031	7.29E-06	*ARF6*
11	56192263	C\A	0.36	-0.135	0.03	7.29E-06	*COX10*
1	165022857	C\T	0.17	-0.144	0.032	7.34E-06	*NOVA1*
24	21009749	C\T	0.21	-0.139	0.031	7.48E-06	*FLVCR2*
1	183771607	C\A	0.28	-0.141	0.031	7.59E-06	*KLHDC2*
18	9493407	A\G	0.38	0.146	0.033	7.71E-06	*TFCP2L1*
2	105649694	C\T	0.11	0.141	0.031	7.79E-06	*IL2*
18	9189622	G\T	0.32	0.153	0.034	7.83E-06	*CLASP1*
15	65996462	G\A	0.22	0.14	0.031	7.94E-06	*CAPN13*
10	14099221	T\C	0.33	0.136	0.03	8.03E-06	*CD177*
1	164809494	A\G	0.13	0.144	0.032	8.32E-06	*NOVA1*
1	164685130	T\C	0.12	0.144	0.032	8.34E-06	*NOVA1*
15	67018469	T\C	0.13	0.143	0.032	8.38E-06	*ALK*
18	9242801	G\A	0.34	0.156	0.035	8.39E-06	*CLASP1*
18	9263983	G\T	0.34	0.156	0.035	8.39E-06	*CLASP1*
18	9275291	A\G	0.34	0.156	0.035	8.39E-06	*CLASP1*
18	9284287	A\G	0.34	0.156	0.035	8.39E-06	*CLASP1*
2	1794459	A\C	0.08	0.138	0.031	8.76E-06	*OMA1*

Table D.8 *Continued on next page*

Table D.8 *Continued from previous page*

chromosome	position	alleles	maf	beta estimate	standard error	p-value	nearest gene
9	74571038	C\T	0.12	-0.134	0.03	8.76E-06	*ZFAT*
30	5917999	G\A	0.44	-0.136	0.031	8.83E-06	*CNST*
15	66836058	C\T	0.18	0.149	0.033	8.85E-06	*ALK*
15	66014620	A\G	0.17	0.139	0.031	8.86E-06	*CAPN13*
9	74557346	T\C	0.11	-0.133	0.03	9.21E-06	*ZFAT*
24	20894432	G\A	0.1	-0.146	0.033	9.31E-06	*BATF*
24	20895972	A\G	0.1	-0.146	0.033	9.31E-06	*BATF*
24	20896831	C\T	0.1	-0.146	0.033	9.31E-06	*BATF*
24	20899682	T\C	0.1	-0.146	0.033	9.31E-06	*BATF*
24	21000833	A\G	0.1	-0.146	0.033	9.31E-06	*FLVCR2*
10	14461597	G\A	0.33	0.133	0.03	9.50E-06	*LYPD3*

Table D.9: Significance values for all ACTH SNPs above suggestive threshold

chromosome	position	alleles	maf	beta estimate	standard error	p-value	nearest gene
22	37596454	C\T	0.23	0.219	0.033	2.14E-11	*SNORD12*
22	37601007	T\C	0.23	0.215	0.033	7.51E-11	*KCNB1*
21	10999900	A\G	0.13	0.223	0.035	1.90E-10	*KIF2A*
22	37297516	A\C	0.23	0.191	0.034	2.32E-08	*ARFGEF2*
21	23724278	T\C	0.18	-0.182	0.033	4.04E-08	*SEPP1*
21	23726763	C\T	0.18	-0.182	0.033	4.04E-08	*SEPP1*
21	23727790	G\A	0.18	-0.182	0.033	4.04E-08	*SEPP1*
21	11050713	G\T	0.32	0.184	0.035	9.77E-08	*KIF2A*
22	37302580	T\C	0.23	0.182	0.034	1.27E-07	*ARFGEF2*
25	14650611	T\C	0.09	0.164	0.032	4.05E-07	*EPB41L4B*
22	37594118	C\T	0.31	0.168	0.033	4.07E-07	*SNORD12*
21	11322078	G\T	0.1	0.174	0.035	6.44E-07	*ZSWIM6*
21	11206546	C\T	0.08	0.18	0.036	6.74E-07	*KIF2A*
21	11306260	A\G	0.08	0.18	0.036	6.74E-07	*ZSWIM6*
21	11318205	C\T	0.08	0.18	0.036	6.74E-07	*ZSWIM6*
21	11347296	C\T	0.08	0.18	0.036	6.74E-07	*ZSWIM6*
21	11370637	T\G	0.08	0.18	0.036	6.74E-07	*ZSWIM6*
21	11392097	A\G	0.08	0.18	0.036	6.74E-07	*ZSWIM6*
21	11323510	C\T	0.1	0.175	0.035	7.05E-07	*ZSWIM6*
21	11115562	C\T	0.08	0.181	0.037	7.57E-07	*KIF2A*
13	24087845	C\T	0.25	0.184	0.037	8.73E-07	*SCNN1B*
21	10410234	C\T	0.13	0.173	0.035	8.80E-07	*IPO11*
3	101439577	G\A	0.26	0.167	0.034	9.52E-07	*GBA3*
21	11135499	G\A	0.08	0.177	0.036	1.01E-06	*KIF2A*
3	101389612	T\C	0.26	0.169	0.035	1.05E-06	*GBA3*
22	37213519	C\A	0.24	0.167	0.034	1.21E-06	*ARFGEF2*
22	37294366	T\C	0.05	0.166	0.034	1.29E-06	*ARFGEF2*
3	101429436	C\T	0.27	0.165	0.034	1.42E-06	*GBA3*
3	101435557	T\C	0.27	0.165	0.034	1.42E-06	*GBA3*
3	101447867	G\A	0.27	0.165	0.034	1.42E-06	*GBA3*
3	101439891	A\G	0.27	0.165	0.034	1.46E-06	*GBA3*
3	101457464	C\A	0.27	0.163	0.034	1.57E-06	*GBA3*
22	37684328	C\T	0.06	0.161	0.034	1.74E-06	*KCNB1*
20	28706883	T\C	0.22	-0.152	0.032	1.79E-06	*OR2H1*
3	101426791	C\T	0.32	0.161	0.034	1.98E-06	*GBA3*
22	38275543	A\G	0.1	0.163	0.035	2.59E-06	*TMEM189*
3	101445648	G\T	0.26	0.162	0.035	2.73E-06	*GBA3*

Table D.9 *Continued on next page*

Table D.9 *Continued from previous page*

chromosome	position	alleles	maf	beta estimate	standard error	p-value	nearest gene
22	37709717	T\G	0.07	0.153	0.033	3.70E-06	*KCNB1*
3	101391352	G\T	0.24	0.16	0.035	3.74E-06	*GBA3*
3	101469263	A\C	0.26	0.16	0.035	3.85E-06	*GBA3*
21	10995125	C\A	0.16	0.166	0.036	4.18E-06	*KIF2A*
22	37595803	C\T	0.46	0.151	0.033	4.23E-06	*SNORD12*
13	24073474	C\A	0.21	0.168	0.037	4.39E-06	*SCNN1B*
21	11307724	C\T	0.27	0.162	0.035	4.50E-06	*ZSWIM6*
22	37696602	C\T	0.08	0.153	0.033	4.51E-06	*KCNB1*
22	38277626	A\G	0.1	0.159	0.035	4.53E-06	*TMEM189*
22	37340019	A\G	0.21	-0.153	0.033	4.60E-06	*ARFGEF2*
22	37509740	T\C	0.27	-0.156	0.034	4.67E-06	*DDX27*
22	37592331	T\C	0.31	0.154	0.034	4.91E-06	*SNORD12*
22	37350157	C\T	0.21	-0.152	0.033	5.24E-06	*ARFGEF2*
22	37683894	C\A	0.06	0.152	0.033	5.33E-06	*KCNB1*
22	37696485	A\G	0.06	0.152	0.033	5.33E-06	*KCNB1*
22	37251815	G\A	0.05	0.154	0.034	5.43E-06	*ARFGEF2*
22	37253365	G\A	0.05	0.154	0.034	5.43E-06	*ARFGEF2*
22	38255100	T\C	0.1	0.155	0.034	5.86E-06	*TMEM189*
21	10979760	G\A	0.28	0.157	0.035	5.92E-06	*KIF2A*
3	101429197	T\C	0.27	0.155	0.034	6.72E-06	*GBA3*
22	40740864	C\A	0.39	0.133	0.03	7.19E-06	*TSHZ2*
22	37327242	C\T	0.05	0.155	0.035	8.25E-06	*ARFGEF2*
22	37401442	G\A	0.06	0.15	0.034	8.35E-06	*CSE1L*
22	37303532	G\A	0.23	-0.151	0.034	8.41E-06	*ARFGEF2*
22	37751615	G\A	0.07	0.148	0.033	8.52E-06	*PTGIS*
22	38232323	T\C	0.1	0.154	0.035	8.89E-06	*TMEM189*
22	37422561	C\A	0.35	-0.149	0.034	9.07E-06	*STAU1*
20	28682224	G\A	0.22	-0.143	0.032	9.36E-06	*UBD*
21	23709379	G\A	0.06	0.145	0.033	9.75E-06	*SEPP1*
21	23736680	C\T	0.06	0.145	0.033	9.75E-06	*SEPP1*
22	37647086	C\T	0.31	-0.151	0.034	9.78E-06	*KCNB1*

Table D.10: Significance values for all leptin SNPs above suggestive threshold

chromosome	position	alleles	maf	beta estimate	standard error	p-value	nearest gene
19	49581025	G\A	0.23	-0.187	0.037	5.42E-07	*CCDC54*
19	48941767	A\G	0.23	-0.182	0.037	1.14E-06	*BBX*
19	48987630	A\G	0.23	-0.18	0.037	1.31E-06	*BBX*
19	49102373	T\G	0.24	-0.18	0.037	1.32E-06	*BBX*
19	48936677	T\C	0.22	-0.179	0.037	1.50E-06	*BBX*
19	48942697	G\A	0.22	-0.179	0.037	1.50E-06	*BBX*
19	49011908	C\T	0.23	-0.178	0.037	1.78E-06	*BBX*
19	49038033	T\C	0.23	-0.178	0.037	1.78E-06	*BBX*
19	49022570	T\C	0.23	-0.177	0.037	1.98E-06	*BBX*
19	49050327	C\T	0.23	-0.177	0.037	1.98E-06	*BBX*
19	49055203	G\A	0.23	-0.177	0.037	1.98E-06	*BBX*
19	49068376	T\C	0.23	-0.176	0.037	1.99E-06	*BBX*
19	49500590	A\G	0.22	-0.178	0.038	2.22E-06	*CCDC54*
19	49444445	A\G	0.23	-0.176	0.037	2.47E-06	*CCDC54*
19	11111059	T\C	0.42	-0.18	0.038	2.56E-06	*RPL22L1*
14	47283907	G\A	0.23	0.181	0.039	2.79E-06	*MEGF10*
19	49601357	T\C	0.23	-0.175	0.037	2.80E-06	*CCDC54*

Table D.10 *Continued on next page*

Table D.10 *Continued from previous page*

chromosome	position	alleles	maf	beta estimate	standard error	p-value	nearest gene
19	49540241	A\C	0.23	-0.175	0.038	3.04E-06	*CCDC54*
19	49485717	T\C	0.22	-0.175	0.038	3.34E-06	*CCDC54*
19	49492656	G\A	0.22	-0.174	0.038	3.85E-06	*CCDC54*
19	49620200	C\A	0.32	-0.168	0.037	4.46E-06	*CCDC54*
1	54004249	G\A	0.21	0.175	0.038	5.43E-06	*CTNNA3*
19	49459767	G\T	0.22	-0.17	0.037	5.63E-06	*CCDC54*
2	51826889	A\G	0.1	0.161	0.036	5.78E-06	*RHOBTB2*
19	49554094	T\C	0.23	-0.17	0.038	5.82E-06	*CCDC54*
3	25187849	C\T	0.18	-0.173	0.038	5.92E-06	*CNTNAP4*
19	49032937	T\C	0.25	-0.167	0.038	9.04E-06	*BBX*
19	49529333	T\G	0.22	-0.167	0.038	9.15E-06	*CCDC54*
6	38093791	G\A	0.13	0.155	0.035	9.32E-06	*KLRC1*

Table D.11: Significance values for all adiponectin SNPs above suggestive threshold

chromosome	position	alleles	maf	beta estimate	standard error	p-value	nearest gene
15	21018128	A\G	0.11	-0.2	0.033	1.85E-09	*ST6GALNAC2*
15	21026780	A\G	0.11	-0.196	0.033	4.38E-09	*ST6GALNAC2*
15	21025772	T\C	0.1	-0.193	0.034	8.67E-09	*ST6GALNAC2*
22	18132778	C\T	0.15	-0.194	0.035	2.83E-08	*SLC23A2*
15	22055134	G\A	0.07	-0.183	0.034	5.81E-08	*RN7SKP218*
15	21431973	A\C	0.11	-0.184	0.034	7.33E-08	*ST6GALNAC2*
15	21500853	C\T	0.12	-0.185	0.035	9.16E-08	*ST6GALNAC2*
15	21475823	G\A	0.14	-0.186	0.035	9.74E-08	*ST6GALNAC2*
15	21178845	A\G	0.38	-0.175	0.033	1.09E-07	*ST6GALNAC2*
15	22054804	G\A	0.09	-0.179	0.034	1.32E-07	*RN7SKP218*
24	24582154	G\A	0.15	-0.171	0.033	1.88E-07	*NRXN3*
15	21999784	C\A	0.07	-0.176	0.034	2.36E-07	*RN7SKP218*
6	68035083	G\A	0.11	0.178	0.035	2.71E-07	*DIP2B*
15	22035163	G\T	0.08	-0.17	0.033	3.61E-07	*RN7SKP218*
15	21500312	T\G	0.18	-0.173	0.034	3.84E-07	*ST6GALNAC2*
22	18140627	T\C	0.15	-0.173	0.034	4.41E-07	*SLC23A2*
24	24582743	T\C	0.1	-0.16	0.032	5.19E-07	*NRXN3*
15	21452496	C\T	0.12	-0.17	0.034	7.06E-07	*ST6GALNAC2*
15	21414572	C\A	0.12	-0.167	0.034	7.33E-07	*ST6GALNAC2*
15	21271186	C\T	0.14	-0.168	0.034	7.50E-07	*ST6GALNAC2*
15	22052424	C\T	0.07	-0.169	0.034	8.06E-07	*RN7SKP218*
2	85771023	G\A	0.17	-0.169	0.034	8.62E-07	*TTC29*
4	37502360	C\T	0.23	-0.165	0.033	8.96E-07	*GNGT1*
22	18212199	G\A	0.06	-0.167	0.034	9.26E-07	*SLC23A2*
15	22026177	C\T	0.08	-0.163	0.033	9.67E-07	*RN7SKP218*
15	22000800	C\T	0.08	-0.165	0.034	1.07E-06	*RN7SKP218*
15	21397014	C\A	0.12	-0.165	0.034	1.15E-06	*ST6GALNAC2*
4	37502144	C\T	0.23	-0.161	0.033	1.31E-06	*GNGT1*
15	21289677	A\G	0.13	-0.163	0.034	1.75E-06	*ST6GALNAC2*
15	21366722	G\T	0.16	-0.161	0.034	1.80E-06	*ST6GALNAC2*
22	18191701	G\A	0.07	-0.164	0.034	1.96E-06	*SLC23A2*
24	24557290	T\C	0.09	-0.155	0.033	2.21E-06	*NRXN3*
22	18181290	C\T	0.06	-0.16	0.034	2.38E-06	*SLC23A2*
20	3585516	G\T	0.35	-0.155	0.033	2.85E-06	*RIPK1*
22	18135640	C\T	0.06	-0.159	0.034	2.86E-06	*SLC23A2*
22	18196266	C\A	0.06	-0.16	0.034	3.14E-06	*SLC23A2*

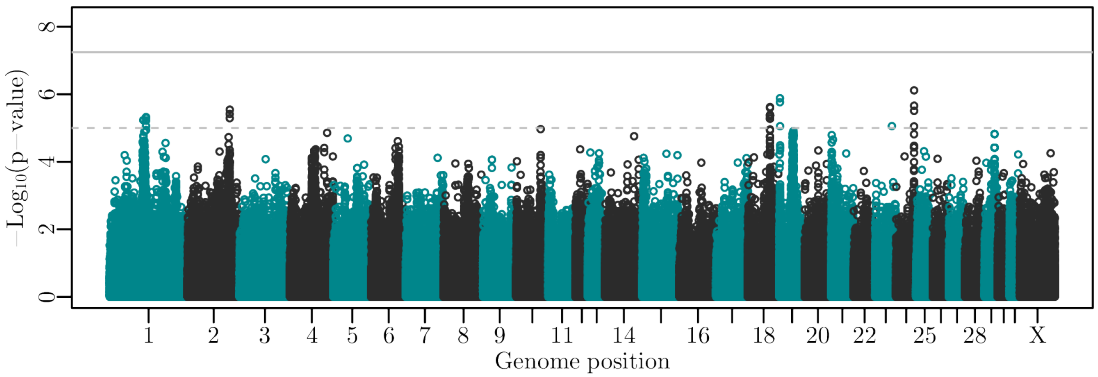
Table D.11 *Continued on next page*

Table D.11 *Continued from previous page*

chromosome	position	alleles	maf	beta estimate	standard error	p-value	nearest gene
15	21631031	G\A	0.06	-0.156	0.033	3.20E-06	*RN7SKP218*
2	17691609	T\C	0.11	-0.161	0.035	3.91E-06	*ZFP69B*
15	21380094	C\A	0.16	-0.155	0.033	3.95E-06	*ST6GALNAC2*
15	21154707	C\T	0.42	-0.155	0.034	4.27E-06	*ST6GALNAC2*
4	37196258	C\T	0.26	-0.155	0.034	4.79E-06	*CALCR*
6	68036518	G\A	0.11	0.159	0.035	4.81E-06	*DIP2B*
4	37412203	C\A	0.44	-0.149	0.033	4.84E-06	*GNGT1*
19	44097058	C\A	0.1	-0.154	0.034	4.86E-06	*GRAMD1C*
22	18154533	C\A	0.06	-0.157	0.034	4.90E-06	*SLC23A2*
15	21281057	C\T	0.22	-0.152	0.033	5.34E-06	*ST6GALNAC2*
15	21150905	C\T	0.39	-0.155	0.034	6.09E-06	*ST6GALNAC2*
15	21387001	G\T	0.14	-0.155	0.034	6.11E-06	*ST6GALNAC2*
15	21331324	T\C	0.16	-0.152	0.034	6.31E-06	*ST6GALNAC2*
24	24540980	A\G	0.09	-0.146	0.033	7.21E-06	*NRXN3*
15	80933450	G\A	0.13	-0.151	0.034	8.17E-06	*TRIB2*
20	3575970	T\C	0.36	-0.15	0.034	8.42E-06	*RIPK1*

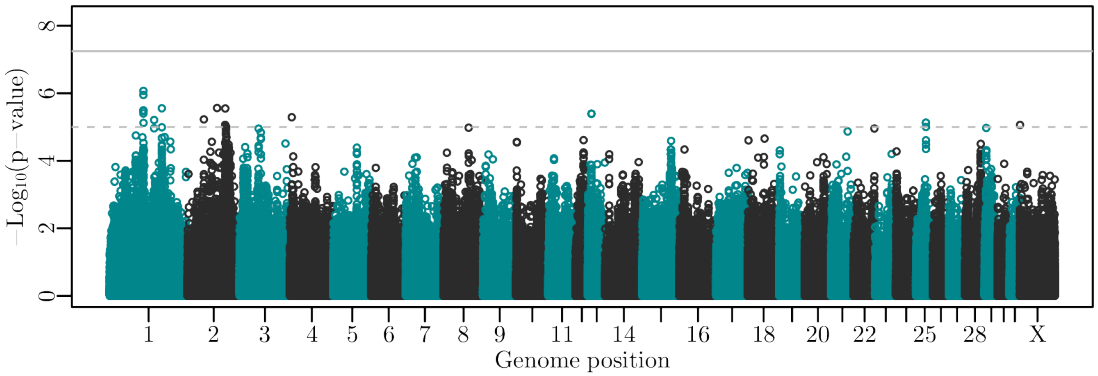
D.2 GWAS results obtained for equine metabolic traits using the standard linear mixed model algorithm provided in the GEMMA software.

Figure D.12: Manhattan plot of GWAS results for neck circumference to height ratio. Bonferonni threshold : p-value$<\mathbf{5.70x10^{-8}}$; suggestive threshold (dashed line): p-value$<\mathbf{1x10^{-5}}$.



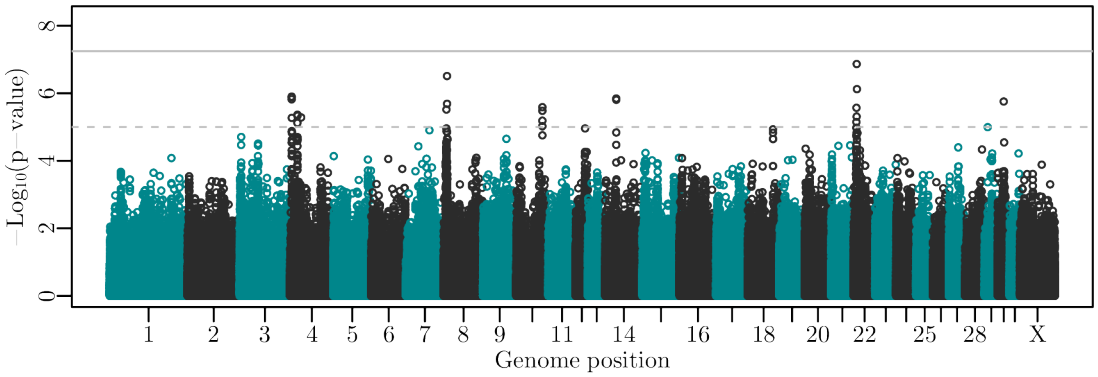
chromosome	position	alleles	maf	beta estimate	standard error	p-value	nearest gene
24	42014088	T\C	0.06	-0.825	0.167	7.67E-07	*EVL*
19	1139669	C\T	0.34	-0.426	0.088	1.32E-06	*VEPH1*
18	59290188	C\T	0.29	0.444	0.094	2.39E-06	*CERKL*
2	96094409	G\A	0.38	0.475	0.101	2.86E-06	*ENO1*
1	85323285	T\C	0.08	0.696	0.152	4.73E-06	*CCSER2*
1	78271748	G\A	0.23	0.459	0.101	5.84E-06	*MAP10*
23	44703897	A\G	0.09	0.658	0.148	8.78E-06	*CAAP1*

Figure D.13: Manhattan plot of GWAS results for girth to height ratio.
Bonferonni threshold : p-value$<\mathbf{5.70x10^{-8}}$; suggestive threshold (dashed line): p-value$<\mathbf{1x10^{-5}}$.



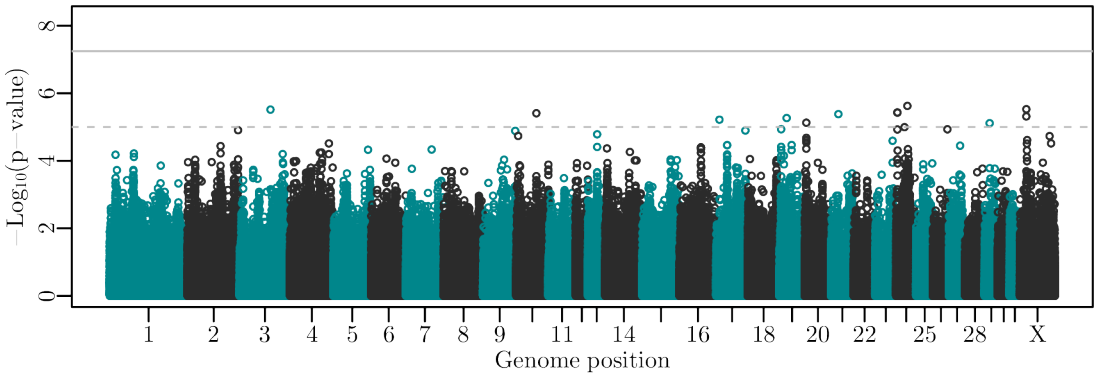
chromosome	position	alleles	maf	beta estimate	standard error	p-value	nearest gene
1	78254375	G\A	0.25	0.48	0.097	8.50E-07	*MAP10*
2	63795958	T\C	0.11	-0.618	0.132	2.74E-06	*GALNTL6*
1	121633670	G\C	0.47	0.413	0.088	2.79E-06	*THSD4*
2	84992300	T\C	0.31	0.418	0.089	2.83E-06	*NR3C2*
13	9371403	T\C	0.29	0.423	0.092	4.05E-06	*SRRM3*
4	3678246	T\C	0.05	-0.816	0.179	5.13E-06	*FGL2*
2	33817044	T\C	0.08	0.673	0.149	5.94E-06	*VWA5B1*
1	103364221	C\T	0.21	0.48	0.106	6.19E-06	*ARRDC4*
25	23370559	C\T	0.07	-0.722	0.161	7.43E-06	*BRINP1*
X	105199	T\G	0.25	-0.457	0.103	8.68E-06	*ZBED1*

Figure D.14: Manhattan plot of GWAS results for fasting glucose.
Bonferonni threshold : p-value$<\mathbf{5.70x10^{-8}}$; suggestive threshold (dashed line): p-value$<\mathbf{1x10^{-5}}$.



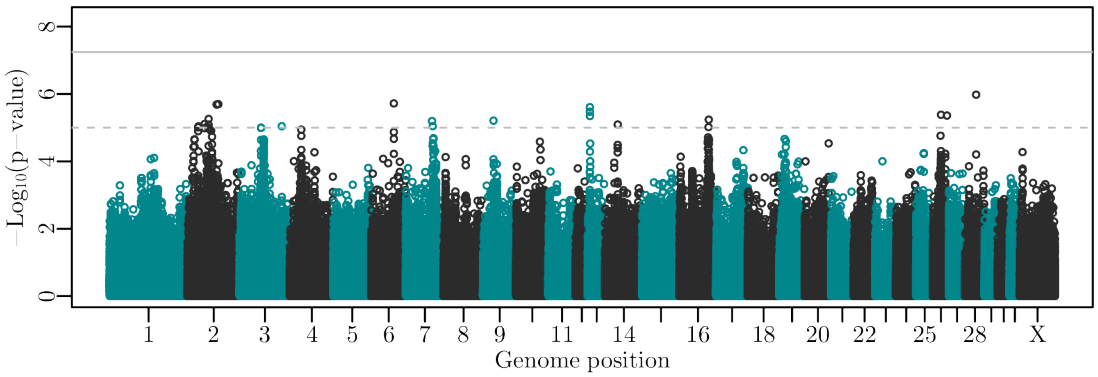
chromosome	position	alleles	maf	beta estimate	standard error	p-value	nearest gene
22	6982098	G\A	0.06	-0.803	0.152	1.36E-07	*BANF2*
8	9312611	T\C	0.31	0.469	0.092	3.10E-07	*CRYBA4*
4	3518141	C\T	0.11	-0.661	0.136	1.25E-06	*CCDC146*
14	26279221	A\G	0.05	-0.885	0.184	1.43E-06	*GLRA1*
30	14742954	A\G	0.26	-0.503	0.105	1.76E-06	*GPATCH2*
10	69248486	G\A	0.11	-0.611	0.13	2.58E-06	*GJA1*
4	18520506	A\G	0.37	-0.412	0.09	4.32E-06	*ABCA13*
4	27797636	A\G	0.29	0.426	0.093	5.21E-06	*PCLO*

Figure D.15: Manhattan plot of GWAS results for fasting insulin.
Bonferonni threshold : p-value$<\mathbf{5.70x10^{-8}}$; suggestive threshold (dashed line): p-value$<\mathbf{1x10^{-5}}$.



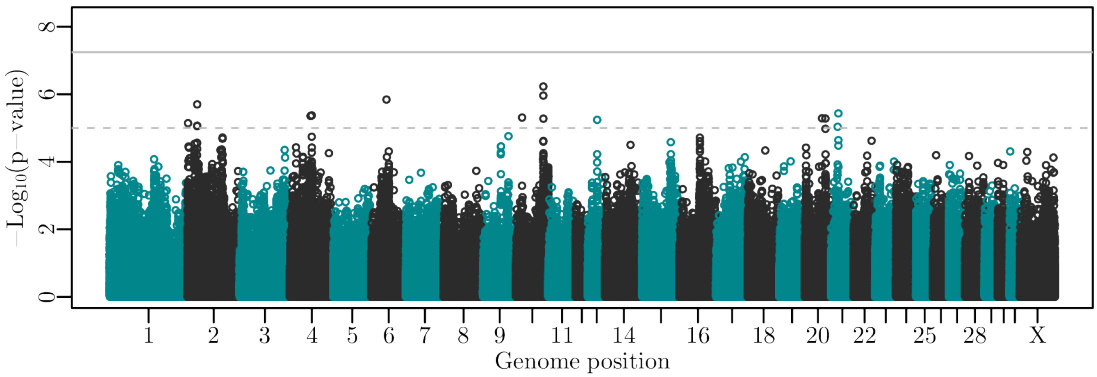
chromosome	position	alleles	maf	beta estimate	standard error	p-value	nearest gene
24	26630151	A\C	0.11	0.646	0.137	2.37E-06	*SEL1L*
X	21928679	G\A	0.38	0.37	0.079	2.97E-06	*MAGEB10*
3	74183327	T\G	0.1	0.645	0.138	3.05E-06	*IGFBP7*
24	3595302	C\T	0.07	0.721	0.156	3.69E-06	*ATG14*
10	53891179	T\C	0.15	0.516	0.112	3.91E-06	*GRIK2*
21	20162158	C\A	0.12	-0.573	0.124	4.14E-06	*ISL1*
19	18465605	G\A	0.19	0.497	0.109	5.41E-06	*MFN1*
17	7858033	C\A	0.13	-0.547	0.121	6.06E-06	*PAN3*
20	4274738	T\C	0.36	0.422	0.094	7.44E-06	*FAM50B*
29	13581131	A\G	0.24	-0.483	0.108	7.70E-06	*ARMC3*

Figure D.16: Manhattan plot of GWAS results for glucose 75 minute post-oral sugar test.
Bonferonni threshold : p-value$<\mathbf{5.70x10^{-8}}$; suggestive threshold (dashed line): p-value$<\mathbf{1x10^{-5}}$.



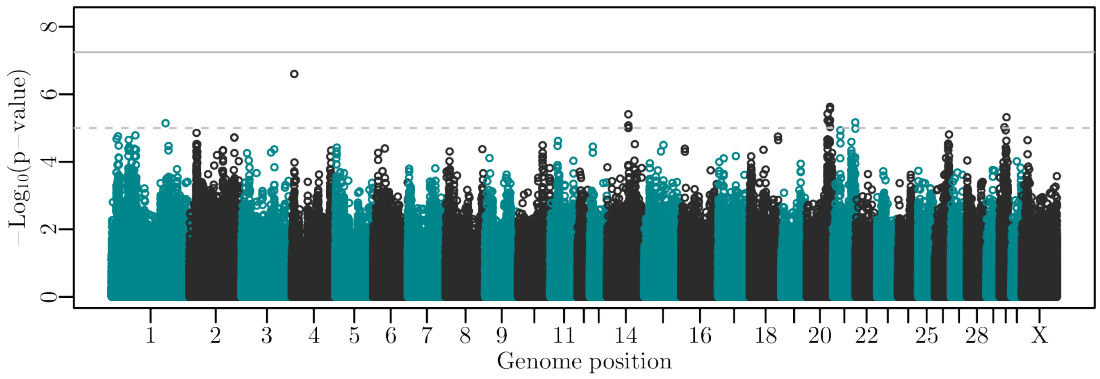
chromosome	position	alleles	maf	beta estimate	standard error	p-value	nearest gene
28	26398304	AG	0.06	0.808	0.165	1.05E-06	*PAH*
6	50244553	CT	0.06	0.845	0.177	1.91E-06	*BCAT1*
2	66782128	TC	0.17	0.553	0.116	2.01E-06	*DDX60*
2	62150475	GT	0.06	0.784	0.165	2.05E-06	*SCRG1*
13	4944861	AG	0.1	0.705	0.15	2.46E-06	*FTSJ2*
26	22041705	GA	0.14	0.555	0.12	4.15E-06	*MIR155*
26	37222496	TC	0.31	0.457	0.099	4.37E-06	*KRTAP10-4*
2	43256184	AG	0.23	0.489	0.108	5.51E-06	*CAMTA1*
16	69351494	TC	0.14	-0.596	0.131	5.79E-06	*TF*
9	24881334	CT	0.07	0.802	0.177	6.19E-06	*TOX*
7	69006360	CA	0.26	-0.428	0.095	6.34E-06	*OR2AT4*
2	35518476	GA	0.06	0.807	0.18	7.78E-06	*IGSF21*
14	29997110	GA	0.24	0.48	0.108	8.09E-06	*SPINK5*
3	100623132	CA	0.36	-0.423	0.095	9.06E-06	*PPARGC1A*
2	25319052	GA	0.3	-0.434	0.098	9.13E-06	*SNRNP40*

Figure D.17: Manhattan plot of GWAS results for insulin 75 minute post-oral sugar test.
Bonferonni threshold : p-value$<\mathbf{5.70x10^{-8}}$; suggestive threshold (dashed line): p-value$<\mathbf{1x10^{-5}}$.



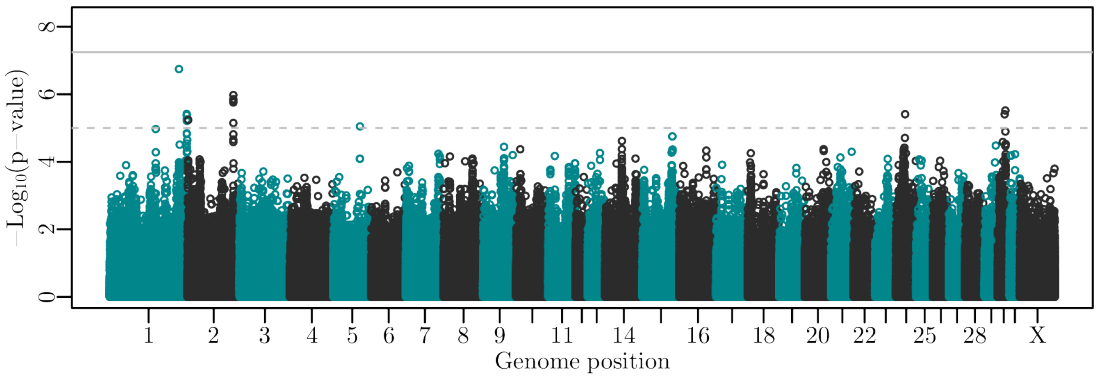
chromosome	position	alleles	maf	beta estimate	standard error	p-value	nearest gene
10	71762350	T\C	0.17	0.573	0.115	5.90E-07	*RNF217*
6	31912608	C\T	0.17	0.59	0.122	1.44E-06	*CCND2*
2	22478751	A\G	0.21	0.498	0.105	2.00E-06	*GJB5*
21	20162158	C\A	0.12	-0.593	0.128	3.70E-06	*ISL1*
4	52134927	T\C	0.36	0.42	0.091	4.24E-06	*MACC1*
10	16809864	A\C	0.34	-0.447	0.098	4.90E-06	*DACT3*
20	42730981	G\T	0.08	0.739	0.162	5.15E-06	*VEGFA*
20	50773321	C\T	0.15	-0.546	0.12	5.24E-06	*GCM1*
13	22421004	A\G	0.06	0.827	0.182	5.73E-06	*ZKSCAN2*
2	1227005	A\C	0.05	0.815	0.182	7.20E-06	*JUN*

Figure D.18: Manhattan plot of GWAS results for trigylcerides.
Bonferonni threshold : p-value$<\mathbf{5.70x10^{-8}}$; suggestive threshold (dashed line): p-value$<\mathbf{1x10^{-5}}$.



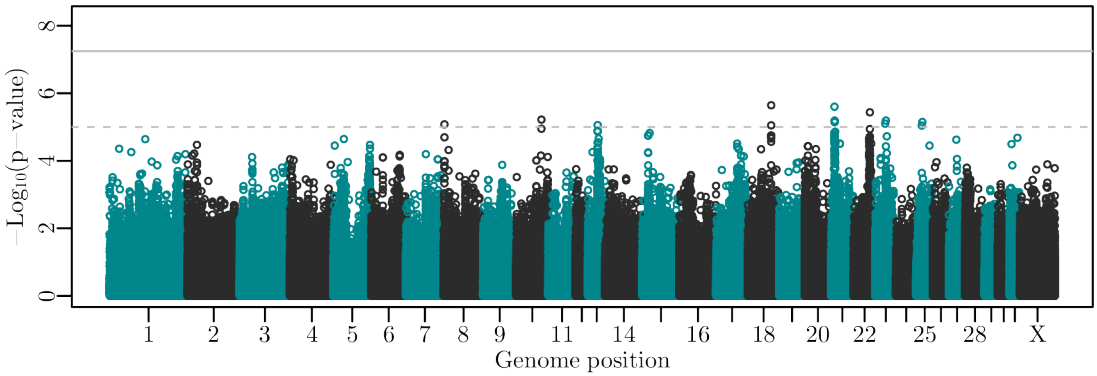
chromosome	position	alleles	maf	beta estimate	standard error	p-value	nearest gene
4	4849780	C\T	0.14	-0.603	0.117	2.51E-07	*LHFPL3*
20	57334198	A\C	0.14	-0.575	0.122	2.37E-06	*NOTCH2NL*
14	52391270	A\G	0.09	-0.664	0.144	3.93E-06	*PRR16*
30	16877429	C\T	0.13	-0.596	0.13	4.81E-06	*KCTD3*
21	56597442	G\A	0.21	-0.457	0.102	6.82E-06	*LPCAT1*
1	125321360	C\T	0.23	0.473	0.105	7.20E-06	*IQCH*
30	11310850	A\G	0.27	0.437	0.099	9.49E-06	*HLX*

Figure D.19: Manhattan plot of GWAS results for non-esterified fatty acids.
Bonferonni threshold : p-value$<\mathbf{5.70x10^{-8}}$; suggestive threshold (dashed line): p-value$<\mathbf{1x10^{-5}}$.



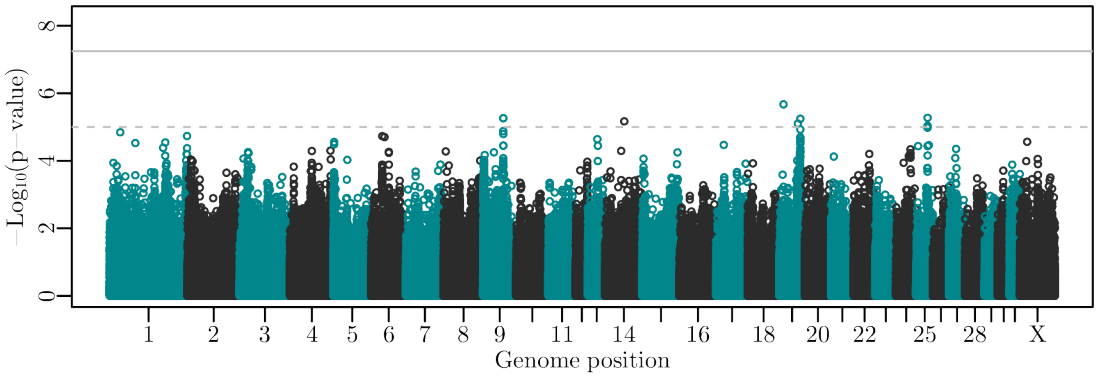
chromosome	position	alleles	maf	beta estimate	standard error	p-value	nearest gene
1	164702224	A\G	0.18	0.558	0.107	1.79E-07	*NOVA1*
2	105698972	A\G	0.1	0.648	0.133	1.05E-06	*ADAD1*
30	19352250	T\G	0.09	0.677	0.145	3.03E-06	*BRINP3*
1	183728003	G\A	0.23	-0.467	0.101	3.84E-06	*KLHDC1*
24	21462649	G\A	0.17	-0.49	0.106	3.91E-06	*TTLL5*
2	1779499	T\C	0.08	0.71	0.157	5.78E-06	*OMA1*
5	70713583	T\G	0.13	-0.545	0.123	8.97E-06	*ABCD3*

Figure D.20: Manhattan plot of GWAS results for adrenocorticotropin hormone. Bonferonni threshold : p-value$<\mathbf{5.70x10^{-8}}$; suggestive threshold (dashed line): p-value$<\mathbf{1x10^{-5}}$.



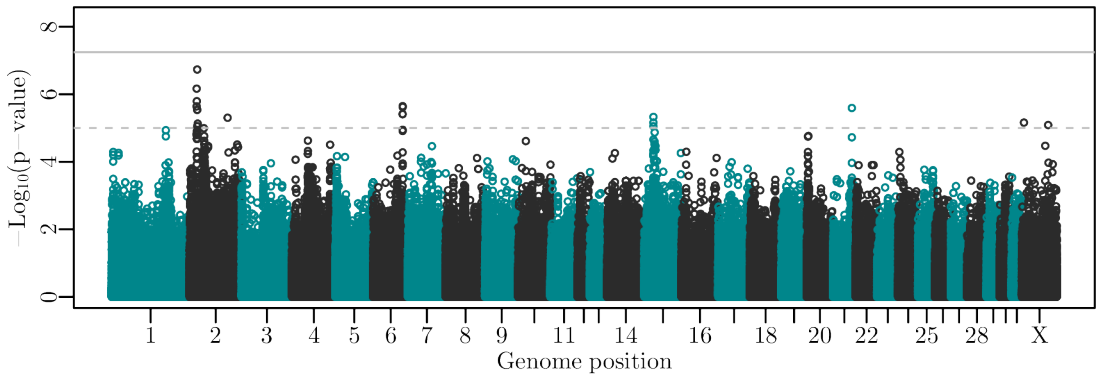
chromosome	position	alleles	maf	beta estimate	standard error	p-value	nearest gene
18	62046771	T\G	0.05	0.842	0.178	2.27E-06	*ZNF804A*
21	10410234	C\T	0.13	0.55	0.117	2.50E-06	*IPO11*
22	37596454	C\T	0.23	0.483	0.104	3.68E-06	*SNORD12*
10	66817465	T\C	0.1	0.624	0.138	6.03E-06	*CEP85L*
23	29059985	A\C	0.14	-0.557	0.124	6.44E-06	*PTPRD*
25	14650611	T\C	0.09	0.617	0.137	7.10E-06	*EPB41L4B*
8	2823191	A\G	0.11	0.569	0.128	8.32E-06	*SLC7A4*
13	24073474	C\A	0.21	0.465	0.105	8.75E-06	*SCNN1B*

Figure D.21: Manhattan plot of GWAS results for leptin.
Bonferonni threshold : p-value$<$**5.70x10^{-8}**; suggestive threshold (dashed line): p-value$<$**1x10^{-5}**.



chromosome	position	alleles	maf	beta estimate	standard error	p-value	nearest gene
19	11111059	T\C	0.42	-0.433	0.091	2.13E-06	*RPL22L1*
25	27382469	G\T	0.28	-0.409	0.09	5.33E-06	*STRBP*
9	50462647	G\A	0.19	0.478	0.105	5.49E-06	*ZFPM2*
19	49581025	G\A	0.23	-0.458	0.101	5.70E-06	*CCDC54*
14	47283907	G\A	0.23	0.449	0.1	6.79E-06	*MEGF10*
19	43369395	G\A	0.22	-0.466	0.104	7.95E-06	*ZBTB20*

Figure D.22: Manhattan plot of GWAS results for adiponectin.
Bonferonni threshold : p-value$<\mathbf{5.70x10^{-8}}$; suggestive threshold (dashed line): p-value$<\mathbf{1x10^{-5}}$.



chromosome	position	alleles	maf	beta estimate	standard error	p-value	nearest gene
2	17691609	T\C	0.11	-0.691	0.133	1.85E-07	*ZFP69B*
6	68035083	G\A	0.11	0.646	0.137	2.25E-06	*DIP2B*
21	48226552	A\G	0.16	0.551	0.117	2.55E-06	*CTNND2*
15	21150905	C\T	0.39	-0.406	0.089	4.68E-06	*ST6GALNAC2*
2	85771023	G\A	0.17	-0.542	0.119	4.96E-06	*TTC29*
X	6005558	G\A	0.06	0.752	0.167	6.89E-06	*GPR143*
X	94742024	T\C	0.29	-0.403	0.09	8.17E-06	*KIAA1210*

Table D.12: Replication of human GWAS candidate genes for metabolic traits using a standard linear mixed model

			Human metabolic trait associations										
Gene	EMS trait	EMS p-value	*adiposity*	*body mass*	*diabetes*	*glucose*	*insulin/IR*	*lipids*	*cardiovascular*	*stroke*	*adiponectin*	*CRP*	*monocytes*
BANF2	GLU	1.36E-07											
NOVA1	NEFA	1.79E-07		X				X	X				
ZFP69B	APN	1.85E-07											
LHFPL3	TG	2.51E-07		X					X	X			
CRYBA4	GLU	3.10E-07											
RNF217	INS OST	5.90E-07					X						
EVL	NH	7.67E-07											
MAP10	GH	8.50E-07											
PAH	GLU OST	1.05E-06											
ADAD1	NEFA	1.05E-06			X								
CCDC146	GLU	1.25E-06											
VEPH1	NH	1.32E-06			X			X					
GLRA1	GLU	1.43E-06								X			
CCND2	INS OST	1.44E-06											
GPATCH2	GLU	1.76E-06			X			X	X				
BCAT1	GLU OST	1.91E-06		X			X			X			
GJB5	INS OST	2.00E-06						X	X				
DDX60	GLU OST	2.01E-06											
SCRG1	GLU OST	2.05E-06				X							
RPL22L1	LEP	2.13E-06							X	X			
DIP2B	APN	2.25E-06											
ZNF804A	ACTH	2.27E-06					X	X	X				
SEL1L	INS	2.37E-06											
NOTCH2NL	TG	2.37E-06											
CERKL	NH	2.39E-06											
FTSJ2	GLU OST	2.46E-06											
IPO11	ACTH	2.50E-06											

Table D.12 *Continued on next page*

Table D.12 *Continued from previous page*

Gene	EMS trait	EMS p-value	Human metabolic trait associations										
			adiposity	*body mass*	*diabetes*	*glucose*	*insulin/IR*	*lipids*	*cardiovascular*	*stroke*	*adiponectin*	*CRP*	*monocytes*
CTNND2	APN	2.55E-06		X	X		X	X	X			X	
GJA1	GLU	2.58E-06							X				
GALNTL6	GH	2.74E-06		X		X	X	X	X				
THSD4	GH	2.79E-06		X				X	X				
NR3C2	GH	2.83E-06						X					
ENO1	NH	2.86E-06											
MAGEB10	INS	2.97E-06											
BRINP3	NEFA	3.03E-06		X		X		X	X	X			X
IGFBP7	INS	3.05E-06							X				
SNORD12	ACTH	3.68E-06											
ATG14	INS	3.69E-06											
ISL1	INS OST	3.70E-06						X	X				
KLHDC1	NEFA	3.84E-06											
GRIK2	INS	3.91E-06		X				X	X	X			
TTLL5	NEFA	3.91E-06											
PRR16	TG	3.93E-06		X		X		X	X				
SRRM3	GH	4.05E-06											
ISL1	INS	4.14E-06						X	X				
MIR155	GLU OST	4.15E-06											
MACC1	INS OST	4.24E-06											
ABCA13	GLU	4.32E-06		X					X				
KRTAP10-4	GLU OST	4.37E-06											
ST6GALNAC2	APN	4.68E-06											
CCSER2	NH	4.73E-06											
KCTD3	TG	4.81E-06											
DACT3	INS OST	4.90E-06											
TTC29	APN	4.96E-06							X				X
FGL2	GH	5.13E-06											

Table D.12 *Continued on next page*

Table D.12 *Continued from previous page*

Gene	EMS trait	EMS p-value	Human metabolic trait associations: *adiposity*	*body mass*	*diabetes*	*glucose*	*insulin/IR*	*lipids*	*cardiovascular*	*stroke*	*adiponectin*	*CRP*	*monocytes*
VEGFA	INS OST	5.15E-06		X	X						X		X
PCLO	GLU	5.21E-06							X				
GCM1	INS OST	5.24E-06							X				
STRBP	LEP	5.33E-06						X					
MFN1	INS	5.41E-06											
ZFPM2	LEP	5.49E-06		X					X				
CAMTA1	GLU OST	5.51E-06		X					X	X			
CCDC54	LEP	5.70E-06		X				X	X				
ZKSCAN2	INS OST	5.73E-06											
OMA1	NEFA	5.78E-06						X					
TF	GLU OST	5.79E-06											
MAP10	NH	5.84E-06											
VWA5B1	GH	5.94E-06		X					X				
CEP85L	ACTH	6.03E-06											
PAN3	INS	6.06E-06								X			
ARRDC4	GH	6.19E-06		X		X		X	X	X			
TOX	GLU OST	6.19E-06		X				X	X		X		X
OR2AT4	GLU OST	6.34E-06											
PTPRD	ACTH	6.44E-06		X	X		X	X	X	X			X
MEGF10	LEP	6.79E-06		X				X	X				
LPCAT1	TG	6.82E-06							X				
GPR143	APN	6.89E-06											
EPB41L4B	ACTH	7.10E-06											
JUN	INS OST	7.20E-06		X				X	X				
IQCH	TG	7.20E-06											
BRINP1	GH	7.43E-06											
FAM50B	INS	7.44E-06											
ARMC3	INS	7.70E-06							X				

Table D.12 *Continued on next page*

Table D.12 *Continued from previous page*

Gene	EMS trait	EMS p-value	Human metabolic trait associations										
			adiposity	*body mass*	*diabetes*	*glucose*	*insulin/IR*	*lipids*	*cardiovascular*	*stroke*	*adiponectin*	*CRP*	*monocytes*
IGSF21	GLU OST	7.78E-06			X			X	X				
ZBTB20	LEP	7.95E-06						X	X				
SPINK5	GLU OST	8.09E-06											
KIAA1210	APN	8.17E-06						X					
SLC7A4	ACTH	8.32E-06											
ZBED1	GH	8.68E-06											
SCNN1B	ACTH	8.75E-06											
CAAP1	NH	8.78E-06											
ABCD3	NEFA	8.97E-06											
PPARGC1A	GLU OST	9.06E-06						X	X	X			X
SNRNP40	GLU OST	9.13E-06											
HLX	TG	9.49E-06						X	X				